THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 54, PART 5 1 September 1942 No. 305

SOME PROBLEMS RELATING TO OPTICAL GLASS

BY W. M. HAMPTON, PH.D., B.Sc.

Inaugural Lecture to the Optical Group. MS. received 30 *March* 1942

§1. INTRODUCTION

WHEN I heard that it had been decided to form an Optical Group of the Physical Society I received the news with considerable pleasure, as I had always felt that the death of the old Optical Society was a great loss, both to science and the optical industry. I feel that I am speaking for the whole optical industry when I say that this new Group will fill a definite need, and I trust that all optical manufacturers will give it their utmost support.

After I had accepted your Secretary's invitation to deliver a lecture on optical glass I found myself in something of a quandary. This was not due to a lack of problems relating to optical glass which could be profitably discussed, but rather to the fact that, owing to the vital character of the optical-glass industry during a war, the restrictions imposed on me by considerations of national safety were such that practically all the points which could be of interest to you were prohibited. It is not too much to say that modern war cannot be fought without optical instruments, and clearly such instruments could not be made unless there was a sufficient supply of optical glass of suitable quality. To mention only a few cases, submarine periscopes, range finders, aerial cameras, all employ optical glass of the highest quality and have, in fact, called for the development of special types of glass to meet their particular needs. Very considerable developments have taken place in the manufacture both of the instrument and the glass, not only in this country but, as you know, in Canada as well. In view of the restrictions imposed by the considerations mentioned earlier, I decided to deal as fully as I was permitted with various specific problems, and for this purpose I considered four main difficulties which face any manufacturer of optical glass under war-time conditions. These are:—

1. The difficulties of supplying glass in sufficient quantity.
2. The difficulty of the supply of raw materials of the necessary quality.
3. The effect of heat treatment on the optical properties of the glass made.
4. The rather more general question of durability.

Under each of these four heads I have taken one particular point and discussed it, and that will form the basis of my paper.

§2. SUPPLY OF GLASS

A Government company, Research Enterprises Limited, was formed during 1940 to erect a factory in Toronto for the manufacture of various instruments, and in some cases the raw materials which were necessary for them. When it was decided, about October 1940, to commence the manufacture of optical glass there, Colonel W. E. Phillips, the President of the company, approached Chance Brothers & Co., Ltd., through the Canadian and British Governments, with the suggestion of co-operation. As a result of that approach I visited Toronto in November 1940 in order to ensure the most rapid development of the new undertaking. The result was that an arrangement was reached whereby the fullest information concerning the manufacturing technique was transferred to the Canadian factory, and certain physicists from Canada came to England and were trained in the art and technique of making optical glass. These physicists remained at our works for some two to three months and were attached to our own personnel, so that they had the fullest opportunity of learning all that we could teach them about the manufacture of optical glass. A number of other people were recruited in this country, some from the Canadian Forces here and some from our own staff, and were given a special intensive course before being sent to Canada for starting up the new process. Further, when manufacture began, one of our technicians went to Canada for some months in order to assist in overcoming the initial difficulties, so that the closest co-operation was maintained and is still being maintained. When it is realized that in November 1940 the optical-glass plant consisted of merely the walls of what was to be a factory, and that the first melting was successfully made in June 1941, and that by the end of 1941 a considerable number of meltings of first-quality glass had been made, it will be realized that, owing very largely to the enthusiasm and drive of the Canadian executives, a very remarkable result has been achieved. There has throughout been the closest co-operation between the Canadians and ourselves, and it can be said that there has been no serious hold-up or difficulty in transferring the manufacture from this country to that.

The factory * occupies about 16 acres, but only a portion of this is devoted to glass manufacture. The pots used are being made by a casting process which is essentially that employed in the Bureau of Standards, namely, the slip-cast method, by which clay treated with a small amount of alkali to give a fluid suspension is poured into a plaster mould. This method, which is well known, has the advantage of being more rapid than the old hand method, and also provides a pot which is considerably drier than that by the old process, and it therefore dries in a relatively short time (figure 1). When required for use, the pot is preheated in one of the batch of pot arches, then transferred to the gas-fired regenerative furnaces and filled with the appropriate batch. Because of the relatively small tonnage employed in an optical glass factory, a mechanical system of weighing is uneconomic, and, therefore, weighing and mixing are done by hand. The scoops and pans used for handling the batch are of stainless steel. All the sand before being used is run through a magnetic separator to avoid any risk of extraneous contamination. The city gas-fired furnaces are of the regenerative type and were built by American companies to R.E.L. designs. The empty

* For fuller details see *Glass Industry* (New York), **22**, Aug. 1941, p. 335.

pots are placed in the furnace by a hydraulic pot carriage, which goes under the name of "Winnie". The rest of the procedure for making, cutting up, moulding, examining, and so on, is an exact copy of that employed at our works in this country, and, in fact, the Canadian factory may be said to be a modernized version of our works. The differences are merely those which are possible because of the existence of an entirely new site and of the possibility of organization from the commencement. The types of glass being made at present are about six, which are the main types in general use for optical instruments. The clays available in Canada have not so far been found appropriate for the more extreme types of glass, but the manufacture of these is visualized, and the continuing co-operation between Canada and Britain will, when necessary, ensure the

Figure 1. Slip-casting of pots at the Canadian factory.

provision of pots of the appropriate resistance to attack to enable all types of glass to be made.

Since the Canadian factory makes not only glass but also instruments, R.E.L.'s customer is in effect itself, whereas in this country the grinding and polishing work is done by separate firms, a method which has both advantages and disadvantages.

In considering the rapidity with which this works was put in hand, it should be remembered that at the time it was being built there was no lighting restriction, so that the factory was built continuously day and night, and there was at that time practically no restriction on the quantity of labour available. These comments are not intended to detract in any way from the achievement of Colonel Phillips and his colleagues, and I can only add that collaboration with them has been simple and complete.

There have been publications from time to time on developments in the manufacture of optical glass. During the last few years, greater developments

have taken place than ever before, due to a sustained programme of research which has been carried out under the aegis of the Admiralty. It is clearly impossible for me to mention in detail many of the developments that have taken place, and I therefore find myself forced to pick out one or two items concerning which information has already been published in this country or elsewhere, and to deal with these only. It follows, therefore, that most of what I have to say today is not new in the sense that it has not previously been published, but it may be of interest to the members of this Society to try and present a connected picture of the various details that I have been able to select.

§ 3. RAW MATERIALS—SAND

Before the war, sand supplies for the manufacture of optical glass were derived practically exclusively from Germany. The sand deposits at Lippe and Hohenbockaer were of the highest quality known, and although substantial stocks had been accumulated or were existent in this country, a serious problem would have arisen had we not been able to find a native source. It had been known for some time (Turner, 1940) that in the Western Highlands at Loch Aline there existed a deposit of sand which appeared to be of fairly high purity as regards iron content. Owing, however, to the fact that it was in a very inaccessible position, the deposit had never been developed before. During the summer of 1940, Messrs. Charles Tennant & Sons, of London, in conjunction with Messrs. Charles Tennant & Co., Ltd., of Glasgow, undertook the development of the site, using the Union Oxide & Chemical Co., Ltd., as the sole agents for England. Considerable quantities have been sent to various glass manufacturers, apart from the requirements for optical glass manufacture. Although the rock on site is very hard, it breaks down readily to yield a granular material suitable in grain for glass melting. It will be realized that the first essential for the manufacture of optical glass is that the sand shall be as free as possible from colouring oxides, and particularly iron. The sand at Loch Aline contains about 0·02 % of Fe_2O_3, which is substantially higher than that of the foreign deposits. Water washing alone, however, reduces this iron content to about 0·012 % Fe_2O_3, which is comparable with Fontainebleau, but it is still substantially higher than the German sand, which used to run about 0·008 %. Intensive experimental work, both at Sheffield under Professor Turner and in our own laboratories, showed that it was possible by mechanical or chemical treatment to remove much of the iron-bearing material, and during the whole of 1941, optical glass in this country was made from Loch Aline sand which had been treated by a Wilfley mechanical separator, with the result that the iron content did not exceed 0·007 %, which is a lower average than had previously been achieved with the foreign material. Recently a new method of treatment has been developed, and sand is now available in sufficient quantities for optical-glass manufacture having an iron content not exceeding about 0·005 %. Thus at the present time the British optical-glass industry has a source of supply of sand which is of substantially higher quality than any which has previously been available. It should be noted, however, that for reasons stated above, and because of the cost of purification, the price of this sand is very substantially higher than that which

was previously available. It is impossible to say at the moment whether it will be economic still to use this sand after the war, but the country is at least independent of foreign sources in case of emergency.

§4. HEAT TREATMENT AND REFRACTIVE INDEX

I will now turn to the question of recent research on the effect of heat treatment on the optical properties of glass, by summarizing briefly work carried out at the British Scientific Instrument Research Association, our own laboratories, and also the literature on the subject. There was an intensive period of research on annealing and its problems between the end of the last war and about 1926. This was all directed towards shortened time-schedule for annealing, and the sole criterion of success was the absence of double refraction. Glass at high temperatures is a viscous liquid, and there is no evidence of any sharp change to a definite solid state as it is cooled to room temperature. The viscosity-temperature curve is continuous, and at room temperature glass can be considered as a liquid of enormous viscosity. In cooling it from a high temperature, therefore, any temperature gradients which exist are, as it were, frozen into the glass, with the result that instead of being optically homogeneous and isotropic it becomes optically inhomogeneous and anisotropic, behaving as a uniaxial crystal. As a result of elastic deformation, variations in refractive index are present, and on examination in polarized light it is found to have different optical properties in different directions. Sufficiently slow cooling reduces these birefringence effects to a negligible amount, and the old test of an annealing schedule was that it should reduce the double refraction to something of the order of $0{\cdot}01\,\lambda$ per cm. It was, however, found on occasion that glass which appeared satisfactory from the point of view of double refraction and which was known to be free from striations still did not give a satisfactory optical image, and it was certainly possible to find two pieces of glass which, while appearing equally satisfactory in polarized light, gave substantially different pictures when examined on an interferometer (figure 2). At the same time it was clear that glass could be used satisfactorily under conditions where, had it been examined in a strain viewer, it would have been seen to be bad. The strain viewer is so sensitive to the effect of temperature changes that it is necessary to maintain a block at a constant temperature for a considerable time and to handle the block with cotton wool or other insulating materials during examination in order to get a satisfactory strain picture. In other words, the warmth of the hand in picking up a block is sufficient to change its appearance rapidly and seriously in the strain viewer, yet this same glass when used, say, in a submarine periscope which might rise suddenly from the temperature of the sea into 20° of frost or bright sunshine, gives an entirely satisfactory performance, so that it is clear its optical performance is not directly related to its strain-viewer appearance. Occasions have been known where lenses have been found unsatisfactory and have been returned for re-annealing, and although from the strain-viewer pattern they looked rather worse after re-annealing than they did before, the optical difficulties had disappeared and the glass was satisfactory. It is clear from all this that there is something other than double refraction which can be affected by the annealing process.

It has long been known that glass which is cooled rapidly has a lower refractive index than glass which has been cooled slowly from the same high temperature. These differences in refractive index are bulk differences, i.e. the same result is found whether the refractive index is measured at a surface, as in the Pulfrich refractometer, or whether it is measured through the bulk of the glass, as on a spectrometer. These differences are relatively large, amounting to, say, six or more units in the third decimal place. It can be shown by simple calculation that the plus and minus variations due to compression or tension in unannealed glass which is stressed to the tensile limit will not exceed more than one or two

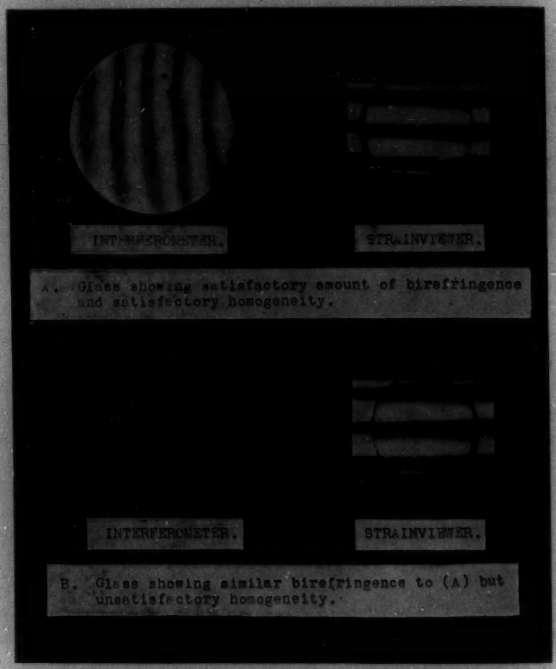

INT-MACRO-SCOPE.
STRAINVIEWER.
A. Glass showing satisfactory amount of birefringence and satisfactory homogeneity.
INTERFEROMETER.
STRAINVIEWER.
B. Glass showing similar birefringence to (A) but unsatisfactory homogeneity.

Figure 2. Interferometer and strain-viewer photographs.

units in the fourth decimal place. These differences due to tension or compression are, therefore, almost completely masked by the general lowering of refractive index, corresponding to cooling rapidly from a high temperature. There has been a great deal of research on this question, as, for instance, on the effect of heating specimens from the same melting, some of which have been chilled and some of which have previously been annealed.

Figure 3 shows the result of one such experiment (Lebedeff, 1926). Small specimens of the glass are heated at a constant rate and are extracted from the furnace at various temperatures. The pieces are allowed to cool rapidly in the

air and the refractive index of the cold glass is measured. It will be seen that up to a temperature of some 500° or 600° c. there is no change in the refractive index of either sample. As the temperature rises, however, there is a marked change in the refractive index, which in the case of the well-annealed samples tends to fall until a temperature of about 800° c. is reached, beyond which it remains constant. In the case of the chilled specimens, the index rises a little and then, at about 800° c., falls to the same curve as the well-annealed glass.

A second series of experiments was to take well-annealed and chilled samples of the same glass and to maintain them at a constant temperature of, say, 550° c. (Twyman and Simeon, 1923; Winter-Klein, 1937). Specimens are extracted from the furnace as before and the refractive index of the cold glass measured. Figure 4 shows the results of such experiments. It will be seen that the well-annealed glass tends to fall in index and the unannealed glass tends to rise in index, and there are indications of an equilibrium index which is a function of

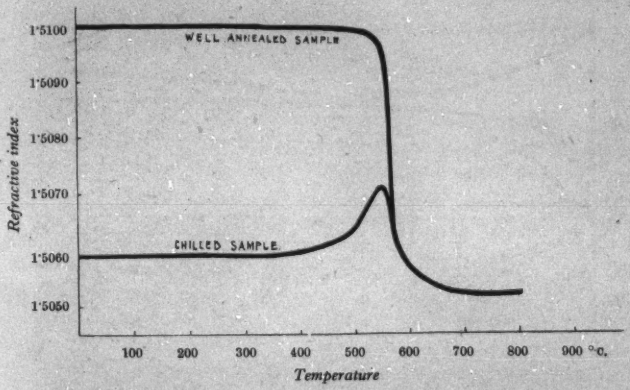


Figure 3. Effect of heating annealed and chilled samples of BSC glass.

the stationary temperature. It is also seen that the equilibrium index decreases as the temperature rises. By determination of equilibrium temperatures in this way, a curve of equilibrium index versus temperature can be plotted, and according to various authorities this curve is approximately linear over the range where, in the first experiments, the index was changing rapidly. The velocity with which the specimens of glass approach the equilibrium index is also a function of temperature, and this velocity increases very rapidly as the temperature rises. It is found that the rate of change of index with time at a constant temperature can be represented by a relation similar to that found by Adams and Williamson (1920) for the rate of change of double refraction. It was shown by the present author (Hampton, 1925/6) that the velocity constant was inversely proportional to the viscosity of the glass. Approximate calculations of the present velocity constant for refractive index, based on such meagre information as is available, suggests that a similar proportionality exists in the case of velocity constants for the approach of refractive index to the equilibrium index, and

although the matter cannot be considered as proved, it is almost certain that this velocity also is inversely proportional to the viscosity of the glass. Since the viscosity of glass falls very rapidly with increase of temperature, the velocity of approach also changes very rapidly, and may vary in the ratio 100 to 1 for a change of, say, 50° c. This fact makes the determination of the {equilibrium index, temperature} curve very difficult, and results at both ends of the annealing range are of doubtful value.

Without at the moment going into the question of what causes this change in the equilibrium index, it is clear that we have here an explanation of the curves in figure 4. Consider the well-annealed sample being heated to a high temperature and samples being removed from the furnace. At the lower temperatures the viscosity is so high that the velocity of approach to the equilibrium index remains what it was at room temperature. Above about 500° c. the

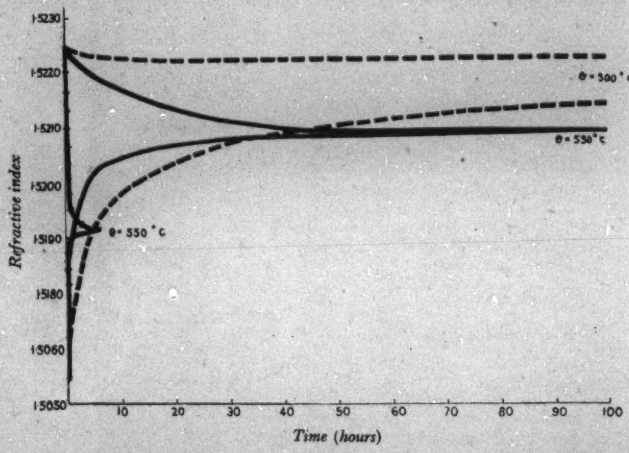


Figure 4. Effect of heating at constant temperatures on index of annealed and chilled samples.

viscosity of the glass is beginning to fall to a reasonable value and the refractive index of the glass turns towards the equilibrium-index line during the heating. It does, in fact, rather overshoot this line, and approaches it asymptotically. It will be noted that the experimental curve shows that the refractive index as measured on cold samples is constant above a certain temperature, whereas extrapolation of the {equilibrium index, temperature} curve suggests that the refractive index should continue to fall. It may be pointed out that the experimental evidence does not disprove the implication of the {equilibrium index, temperature} curve. The velocity of approach to equilibrium at these high temperature is so large that even during the high rate of cooling which is implied by the word " chilling " it is probable that the glass index can readjust itself to the equilibrium value corresponding to the temperature. Thus it is to be expected that the refractive index of a sample chilled from, say, 1000° c. will follow the equilibrium line for some distance until the viscosity becomes too

high for this instantaneous adjustment to take place. The refractive index measured on the chilled sample, therefore, taken at 1000° c., would be recorded as the index corresponding to some lower temperature, and it follows also that, for a given rate of chilling, this recorded index will be independent of the temperature from which the sample was chilled, provided this temperature is sufficiently high. It may be shown experimentally that the limiting index is lowered with increase in the chilling rate.

Various experimenters, notably Lebedeff (1926) and Klein (1937) have assumed that these curves have indicated the existence of two forms of glass which, for want of a better term, they have called the α and β forms by analogy with the silica transformation. They say that glass below about 500° c. exists in the α form and above about 700° or 800° c. in the β form, and between those temperatures there is an equilibrium between the α and β forms. There are three facts which seem to me to throw grave doubt on this supposition:— (1) That the lower value of index reached can be shown to be a function of the rate of chilling during the experiment; (2) that a pure boric-oxide glass free from silica shows the same sort of effect (Turner, 1923); and (3) that the viscosity-temperature curve does not show any marked changes of direction as the temperature of the glass is raised.

It seems to me more reasonable to suppose that glass at very high temperatures consists of an intimate solution of simple silicates which on cooling tend to form aggregations into larger molecules. The important practical point is, of course, that the refractive index of a given glass depends on the effective temperature which is frozen in during the annealing operation. It is clear that the slower the cooling, the higher will be the refractive index.

These results have an important bearing on the annealing of slabs of glass. If a piece of glass is uniform in temperature while being held prior to the annealing operation, the only index changes introduced into it will be the plus or minus variations due to the temperature gradient during cooling, and these will probably be small, perhaps not more than a few units in the fifth decimal place. If a small piece of glass is held at, say, 550° c. until the index has approached the equilibrium and is then cooled slowly from this temperature, it will have a general refractive index slightly higher than the equilibrium index for 550° c., together with the plus or minus variations due to the gradient. If it is held at some other temperature, say 540° c., it will have a final index slightly higher than the equilibrium index for 540° c., together with the plus or minus variations, but the two specimens of glass will have different refractive indices, although their double-refraction pattern may be the same. If, therefore, both these temperatures exist during the annealing operation, that is, if a temperature difference exists during the soaking period, then different parts of the finished plate will have different refractive indices. Without the most elaborate precautions it is very difficult to get any sort of annealing kiln where the temperature differences are so small as not to give a risk of departures across a slab of optical glass of the order of 10° c. Much experimental work, therefore, has been directed towards the provision of annealing equipment where the temperature differences are negligibly small, and kilns have been designed and are in use where the temperature difference over a length of 2 feet or more does not exceed 3 or 4° c. These

kilns are capable of being cooled at a much higher rate than the old type, while still yielding glass which is of the highest standard when viewed on the interferometer, although the strain-viewer pattern may not be appreciably different from that given by less uniform kilns. Experience has shown that the glass which is satisfactory on the interferometer, due to the absolute uniformity of index, is capable of being used for the very highest types of optical instruments, and it is becoming generally accepted in the optical industry that the interferometer, and not the strain-viewer, is the instrument to be used. It is now a routine operation for samples from all kinds of optical glass to be examined on the interferometer, and for the strain viewer to be used to a considerably less extent.

There was for many years an objection, which was based on painful experience, that the highest quality prisms could only be obtained by sawing from large blocks, and not by moulding. Since moulding is much more economical in the use of optical glass than cutting up large pieces, recent developments in annealing, which have proved beyond doubt that moulded prisms can give results equal in quality to those produced by any other method, have been of the utmost importance in amplifying the supply under the present emergency conditions.

§ 5. DURABILITY

Another question which I would like to discuss briefly is that of the determination of the durability of optical glass. The object of such a test is to obtain a measure of the resistance of a polished surface of glass to atmospheric agents, since this surface in an optical element is subject in use to exposure to the varying conditions of humidity and the possible chemical attack of atmospheric impurities. The varying humidity of the atmosphere results in corresponding variations in the water content of the glass surface, accompanied by alternations of deposition and evaporation of moisture. These alternations result in the separation of soluble constituents from the glass and the production of a surface film. Compared with its optical properties, the durability of optical glass is of secondary importance, but it is essential that in obtaining any desired optical properties, either by developing a new type of glass or modifying an existing type, the final choice of batch should be decided by the durability. It is obviously important, therefore, that the manufacturer should be able to determine fairly rapidly the effect on durability of variations in batch composition, and it is also desirable that the user should be furnished with some indication of the probable performance under service conditions. Whereas with well established types, observations on glasses which have been submitted to prolonged exposure to atmospheric conditions are fairly well known, such a method of testing is clearly impracticable as a means of forecasting what a new glass will do. In addition, owing to variations in conditions during actual exposure, such information as is available is of a statistical character. Several laboratory methods have been put forward for determining the durability, but such accelerated attacks as a rule differ markedly from those to which the glass is exposed in use. The results obtained by such methods are sensitive to slight variations in experimental conditions, and in consequence often fail to give agreement even among themselves. In

addition, in the few cases where comparison has been possible with the results obtained under natural but severe conditions, there appears to be little correlation. The tests which have been employed and discussed in the literature are the following:—

Chemical resistance tests (see, for example, Taylor, 1936)

This method is used to determine the chemical resistance of glasses to chemical reagents. The glass is usually in the form of a powder graded between moderately close limits and the solubility of the powder in the reagents is determined. This test is very drastic in character compared with the weak chemical action of the atmosphere, and in general there is little relation between the results given by it and those found in use.

Autoclave test (see, for example, Baillie, 1922)

In this test a polished glass surface is subjected to the action of superheated steam. The rate of corrosion is far more serious than is met with in use, and the conditions of humidity to which the glass is subjected are quite different from those in practice. The test does not as a rule include the variation of temperature conditions, which forms one of the most prominent features in practical use.

Dimming test (see, for example, Elsden, Roberts and Jones, 1919)

This consists in subjecting the surfaces of the glass to the action of air saturated with water vapour at some uniform temperature, usually 80° c., for a definite period. This test does not allow for the humidity variation already mentioned, and very long periods are necessary in order to distinguish between good glasses.

Iodoeosin test (see Mylius, 1910 and 1913)

Results based on the Mylius iodoeosin test are quoted in certain optical-glass catalogues to specify the durability of the various types. A quantitative measurement is usually made of the amount of free alkali in a freshly fractured surface which is available for combination with iodoeosin. The surface is then subjected to exposure to a moist atmosphere at 18° c. for seven days, and the change in the amount of free alkali is determined. Thus, increase in the amount of free alkali is taken as lack of stability, and decrease of free alkali is associated with stability. While it is accepted that the test is probably reliable for glasses containing appreciable amounts of alkali, it is misleading for those types containing lead oxide, zinc oxide and boric oxide.

Electrical test (see B.I.R.A. reports R. 19 and R. 26)

The British Scientific Instrument Research Association devised a test for the durability of optical glass surfaces in which the electrical conductivity and changes of conductivity under accelerated weathering conditions are measured. As with the iodoeosin test, this is only reliable for certain types.

Some recent work on the durability of optical glass has been published in America (Jones, 1941). Observations were made on the behaviour of polished glass surfaces exposed to ordinary service conditions, and two types of surface

change due to weathering were noted. Glasses which had been subjected to a humid atmosphere became covered with a hazy film of soluble alkaline salts, and this process was called " dimming ". On the other hand, contact with condensed moisture led to the separation of lead or barium oxides and the formation of a silica-rich film. The film exhibited interference colours and the process was called " staining ". Various accelerated durability tests were investigated, and it was concluded that two tests were required: one to determine the tendency to dimming and the other to determine the tendency to staining. For the dimming test, exposure to a saturated atmosphere for 28 days at a constant temperature of 50° c. was employed. The best test for staining was found to be that developed by Berger (1936) using nitric-acid solution of known strength at 25° c. and observing the time required for the formation of a surface film showing a dark blue interference colour.

The test which I propose to describe now was developed in this country over a period of some years, and, in fact, was being used as a routine method before the recent American work was published.

In attempting to devise a new test for the determination of durability it was laid down that the method should be capable of discriminating between slight differences in durability, and the variables should be capable of accurate control; it was also considered essential that the results should be comparable with those that are obtained under natural conditions by exposing specimens of the same glasses to the action of the atmosphere in the tropics for periods of one to two years. In practice, the instruments are in many cases subjected to a relatively high temperature during daylight and a low temperature during the night. The result is that moisture condenses on the glass at night and is evaporated off during the day. An apparatus was therefore constructed in which cycles of temperature were introduced, and the name *thermodyne* was coined to describe it, in contradistinction to the thermostat.

The thermodyne method of determining the durability of glass surfaces consists essentially of submitting either freshly broken or optically polished surfaces (sealed off in a flask together with a quantity of water) to a series of temperature cycles from room temperature to 60° c. in an atmosphere of air saturated with water vapour for a period of 12 days. Each cycle is of two hours' duration. The attacked surfaces are subsequently examined visually and photographed for record purposes. The type of attack which results is characteristic of the type of glass. The degree of attack is recorded by an arbitrary number, comparison always being made with the attack on specimens of similar standard type tested simultaneously. Control is possible on the following points:—(*a*) Upper and lower temperature limits and the duration of the cycle; (*b*) humidity of the atmosphere in contact with the specimens; (*c*) composition of the atmosphere in contact with the specimens; (*d*) condition of the test surfaces.

The results obtained with the test were compared systematically with the results which had been obtained by exposure of the same glasses to tropical atmospheres. Without going into details of the experimental procedure, it may be said that this method of test has given valuable results.

It may be pointed out that both the conditions for producing the two types

of weathering noted by the American experimenters are embodied in our method, and while it is recognized that no single test is available for determining the durability of the wide range of optical glasses manufactured, comparison of the results of the thermodyne test with the results of natural-exposure tests shows that the accelerated test is fairly reliable.

In typical tests on binocular prisms returned from service, the specimens appear in the same relative order when ranged according to their state after this test or after natural exposure; moreover, the two specimens which were badly stained naturally show heavy staining in the accelerated test.

It may be said generally that the thermodyne attack is not exactly the same as the natural attack. In fact, the thermodyne attack is rather more severe than that resulting from years of exposure in the tropics, but the more severe nature of the accelerated attack accentuates differences in durability. At the same time the test is not so severe as some of those already described, and it is perhaps most valuable in assessing the durability when modifications of an old type of glass are being developed.

It is possible that with the development of non-reflecting surface films, this question of glass durability will become of less importance, and it is worth noting that Jones and Homer (1941) have recently used the method employed for their accelerated weathering test to produce silica-rich films which have later been stabilized by baking to form a durable film.

§6. CONCLUSION

I do not make any suggestion that the four problems I have dealt with are the only ones with which we are concerned, nor even that they are the most important ones; they were simply taken as a means of indicating to you the type of problem that has to be met in the manufacture of optical glass. I feel that a Society such as this will give opportunities for closer collaboration between makers and users of glass, and I would like to point out that while there are many lines on which development of new types of optical glass may proceed, it is only in a few of these that any useful result is likely to be achieved. In order to decide which of these calls for the closest collaboration between lens designers and the makers of optical glass, it would be of interest and value to us as manufacturers if the representatives of the various lens-designing and lens-making firms could give fairly precise indications as to the type of {refractive index, dispersion} relation that would be of most value to them in designing new systems. At any time, there is a definite restriction on the amount of research which can be undertaken, and we do not wish to extend our exploration over parts of the problem which do not seem likely to lead to fairly rapid results. If, therefore, either through this Society or otherwise, it could be so arranged that the users could give precise figures of index and dispersion that would be of value, much unnecessary work could be avoided. Certain optical firms have already indicated to us the types of glass which they think will be needed in the future, and we hope that other firms will follow suit.

REFERENCES

ADAMS, L. and WILLIAMSON, E., 1920. *J. Franklin Inst.* **190**, 597.
BAILLIE, W. L., 1922. *J. Soc. Glass Tech.* **6**, 279.
BERGER, E., 1936. *J. Soc. Glass Tech.* **20**, 257.
ELSDEN, A. V., ROBERTS, D. and JONES, H. S., 1919. *J. Soc. Glass Tech.* **3**, 52.
HAMPTON, W. M., 1925–6. *Trans. Opt. Soc.* **27**, 163.
JONES, F. L., 1941. *J. Amer. Cer. Soc.* **24**, 119.
JONES, F. L. and HOMER, H. G., 1941. *J. Opt. Soc. Amer.* **31**, 34.
LEBEDEFF, A. A., 1926. *Rev. d'Optique*, no. 1, 1.
MYLIUS, F., 1913. *Silikat Z.* **1**, 2, 25, 45 ; 1910. *Z. Anorg. Chem.* **67**, 200.
TAYLOR, W. C., 1936. *J. Soc. Glass Tech.* **20**, 405.
TURNER, W. E. S., 1923. *J. Soc. Glass Tech.* **7**, 139.
TURNER, W. E. S., 1940. *J. Soc. Glass Tech.* **24**, 197.
TWYMAN, F. and SIMEON, F., 1923. *J. Soc. Glass Tech.* **7**, 199.
WINTER-KLEIN, N., 1937. *Rev. d'Optique*, no. 11, 361.

DISCUSSION

Dr. SPENCER JONES. I am very interested in what Dr. Hampton has said about the thermodyne test for dimming. In the last war I was much concerned with the trouble from dimming of glass surfaces in optical instruments used under tropical conditions. I used both the autoclave and iodoeosin tests and found that there was a close correlation between the results of these tests and the liability to dimming in hot moist climates, though, as Dr. Hampton has mentioned, the tests were not in every respect satisfactory. It was found that, for the same type of glass, the trouble from dimming was dependent both upon the perfection of polish of the surface and upon the care taken in cleaning it. Dr. Hampton's tests are made mainly upon freshly fractured unpolished surfaces, and I should like to ask him whether the thermodyne tests show that the dimming is dependent both upon the perfection of polish of the surface and upon whether or not it is chemically clean.

AUTHOR's reply. It should be pointed out that with the thermodyne test, it is possible to determine the liability to staining as well as the liability to dimming. Both the autoclave and the iodoeosin tests give misleading results for glasses which stain badly under natural conditions, and our experience with the thermodyne test shows that it gives more reliable results for such glasses, particularly for dense barium crowns.

With regard to the condition of the test surfaces, a considerable amount of work was done on this subject during the development of the test. It was found, for example, that unless the surfaces were chemically clean, there was considerable variation in the results. The criterion of cleanliness adopted for the test is that the surface can be uniformly wetted by distilled water.

It has also been found that for certain types of glass the results depend on the perfection of polish of the test surface. Thus, surfaces polished by an inexperienced polisher appeared to be much less durable than those polished by an experienced polisher, and the results for freshly fractured surfaces agreed with those obtained for surfaces polished by the experienced polisher. This result is in agreement with the conclusions reached by Lord Rayleigh (*Proc. Roy. Soc.* A, 1937, **160**, 507) that the refractive index of a polished glass surface depends on the method of polishing, and that for commercial polishing the refractive index of the polished surface approximates to that of the body of the glass.

For this reason, freshly fractured surfaces are used whenever possible in the thermodyne test, and in such a case the only cleansing necessary is the removal of glass débris which arises from the fracture.

THE MEASUREMENT OF RADIATION FOR MEDICAL PURPOSES

By W. V. MAYNEORD,
The Royal Cancer Hospital (Free)

Lecture delivered 27 *February* 1942

ABSTRACT. The measurement of radiation for medical purposes is primarily concerned with the measurement of absorbed energy either at a given element of mass within the irradiated tissues or throughout the body as a whole.

The ionization method of measuring x rays and gamma rays is briefly described in the paper and the fundamental nature of the röntgen is discussed. The distribution of radiation throughout a body subjected to x or gamma irradiation is next considered, and examples given of the variation of dosage-rate in typical cases along the axis of a beam sent into soft tissues. This leads to a study of " isodose surfaces " and some simple geometrical methods of obtaining them.

The relationships between the dose in röntgens and the energy absorbed per unit mass of material or the energy flux per square centimetre of cross-section of the beam are indicated.

The second problem of total energy absorbed is then attacked by integrating the product of the dose and element of mass throughout the whole mass of interest, thus obtaining the so-called " integral dose ". Again, simple geometrical and analytical methods of deducing the values of the quantities involved are described. The significance of total energy absorption in many problems (as, for example, that of protection against stray radiation) are indicated. Finally, the total amounts of energy involved in the therapeutic use of ultra-violet and infra-red radiations are compared with those of x rays and gamma rays, and a plea is entered for the correlation of all these radiations on an energy basis.

§1. INTRODUCTION

The methods of measurement of radiation for any purpose whatsoever must in the end be determined by the fact that we are concerned with a form of energy. Nevertheless the specific purpose to which a radiation is to be put often decides which of a number of quantities it is of greatest importance to investigate. Left to a severely physical investigation of a beam of x rays or gamma rays, we might be tempted to concentrate on the measurement of its intensity, but if the beam were intended for use in radiation therapy our efforts would then be to some extent misdirected.

The whole of the study of the biological and medical use of radiation is dominated by the rule that absorbed energy alone is effective, so that we are forced in considering the physical aspects of the use of radiations for these purposes to look very closely into the measurement of this quantity. This study of absorbed energy, either at any point in the tissues, or throughout the body as a whole, forms the subject matter of this paper.

We may commence with a very short survey of the problem in relation to the ionization system of measurement of x or gamma rays in common use and embodied in the unit known as the röntgen. This leads to an attempt to estimate

the total energy flowing into the body or absorbed by it from beams of x and gamma rays, and eventually to link these values to measurements of ultra-violet and infra-red radiations also used for therapeutic purposes. Our survey is often rough and inexact, but the time seems ripe for an attempt to sketch out the ground, perhaps in this way indicating where detailed examination and experiment may now be most profitable.

§ 2. THE IONIZATION METHOD OF MEASUREMENT OF X RAYS

Among the various physical or chemical effects of x rays, the ionization produced by them in air stands out pre-eminently as the most promising basis on which to construct a system of measurement for medical use. This suitability depends upon many factors, three of which may be noted. First, the measurement of an ionization current may be carried out with fair accuracy over a very wide range of values. Indeed, the measurement itself is often easy compared with the decision under what conditions to measure and what the current means when we have measured it. Secondly, the energy required to produce a pair of ions in air is nearly independent of the speed of the electrons producing it (Eisl, 1929), and so independent of the wave-length of radiation setting the electron in motion. This, together with the fact that the effective atomic number of air is often approximately the same as that of soft animal tissues (Mayneord, 1938), means that the ionization observed in air may be made a suitable basis for the estimation of the energy absorbed by tissues over a wide wave-length region stretching from perhaps 0·01 to 2·0 A. This so-called " independence of wave-length " is very valuable in medical radiology, where the rival, and often conflicting, claims about the superiority of different wave-lengths are keenly debated.

The conditions of measurement thought to be necessary and sufficient to obtain significant ionization currents may be seen in the original definition of the röntgen as adopted by the International Congress of Radiology in 1928.

§ 3. DEFINITION OF THE RÖNTGEN (1928)

" The röntgen is the quantity of x radiation which, when the secondary electrons are fully utilized and the wall effect of the chamber is avoided, produces in 1 c.c. of atmospheric air at 0° c. and 76 cm. of mercury pressure such a degree of conductivity that 1 electrostatic unit of charge is measured at saturation current."

Immediately following the adoption of this unit, experimental and theoretical work was commenced which was intended to extend its use to x rays produced at very high voltages and to gamma rays, since, apart from the theoretical interest of the problem, this was a pressing everyday need. At first considerable difficulty was encountered owing principally to the very great ranges of the secondary electrons set free in air at normal pressures, and the difficulties of eliminating secondary radiations, but eventually the results of measurements with air-wall and parallel-plate chambers were reconciled, and the problem now seems to be fairly satisfactorily solved to the accuracy required for medical use (Gray, 1937; Mayneord and Roberts, 1937; Kaye, Binks and Perry, 1939). It may be added that the research often appeared as the determination of the dose in röntgens

received in one hour at a point one centimetre from a point source of radium filtered with 0·5 mm. platinum. Our own determination of this quantity with air-wall chambers led to the value 8·3 r., in a fair agreement with more recent results obtained by Taylor and Singer (1940) in measurements with a high-pressure parallel-plate chamber which give 8·16 r. as the final value, and probably represent the most accurate determination of this constant to date.

It was not until 1937 that the röntgen was provisionally adopted for gamma rays as well as x rays, a revised definition being at the same time promulgated. This new definition runs as follows:—"The röntgen shall be the quantity of x or gamma radiation such that the associated corpuscular emission per 0·001293 gramme of air produces, in air, ions carrying 1 electrostatic unit of quantity of electricity of either sign."

This new definition aimed at improving the rather clumsy form of the original, extending its use to gamma rays, and also at eliminating from that definition the statement of experimental conditions which gave it the form of the description of a standard rather than that of the definition of a unit. The fundamental idea, namely, that we take all the electrons set free by the radiation from a certain mass of air, utilize as much as possible of their energy in producing ionization in air and measure the number of electrostatic units of charge thus liberated, appears much more clearly in the new definition.

§4. DISTRIBUTION MEASUREMENTS

We now turn to a second aspect of the measurement of x or gamma rays in medicine, namely, the results of measurement throughout a scattering mass. These measurements are undertaken so as to find out how the energy sent into such a mass is distributed, and so, if we can, deduce the energy absorbed per unit mass at each point throughout the absorber, believing that this quantity is of fundamental importance in determining the biological effects of the radiation at each point.

The measurements are normally carried out by substituting for the human body a tank of water throughout which measurements of dose or dosage-rate (that is röntgens per unit time) are made. Owing to the wide range of radiations of interest, different experimental arrangements have naturally to be employed, for we have to solve the problem at low kilovoltages for radiations which penetrate little more than a fraction of a millimetre, while for very high exciting potentials we have radiations which only fall to half their surface dosage-rate at a depth of some 10 cm. to 15 cm. Clearly very different ionization chambers and accessories are required in these two cases.

As an illustration of the kind of results obtained in the normal therapeutic regions of 200 kv. to 2000 kv., we may make use of a survey (Mayneord and Lamerton, 1941) which sets out some of this information.

Considering first of all the change in dosage-rate as we enter the medium along the axis of the beam sent perpendicularly into it, we find that the apparent penetration depends upon three main factors, namely, average wave-length of the radiation; focal distance, that is the distance from source of x rays to the surface of the absorber, and finally to a very great extent on the area of cross-section of the beam which in its turn fixes the mass or volume from which

scattered radiation is received by a point on the axis. These factors were analysed in our survey and tables drawn up from which the doses at a depth may be read for a large number of conditions. Typical examples of the penetrations observed are given in figure 1. Detailed mathematical analysis of the problem has not yet been successfully undertaken, though by the use of the Klein-Nishina scattering formulae, coupled with simplifying assumptions, the main features may be accounted for. The problem is, however, a very complex and difficult one.

Of more interest to us at the moment is the complete distribution of radiation

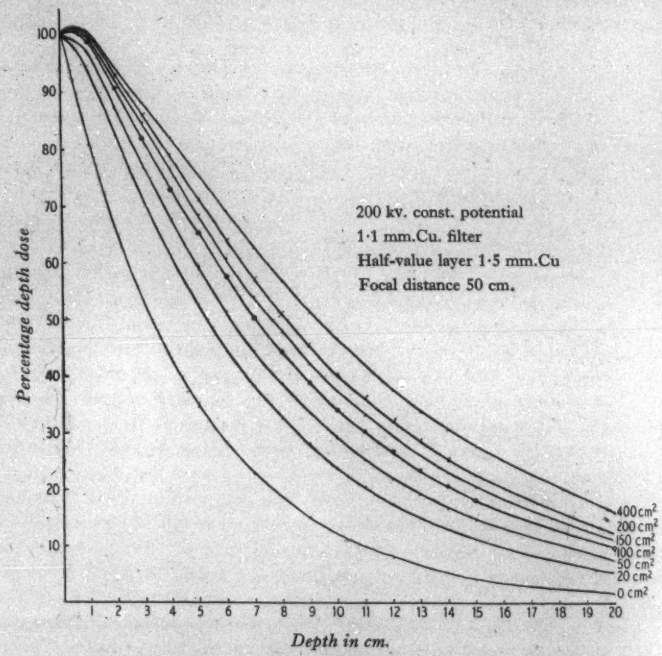


Figure 1. Variation of the dose with depth for a fixed quality of radiation but different field sizes.

in any plane containing the axis of the beam, but of course the experimental measurement is now much more prolonged, so that correspondingly less information is available. However, during the last few years data have been collected over a very wide range of conditions varying from 45 kv. and 2 cm. focal skin distance (Lamerton, 1940) to the gamma rays of radium filtered with 1 cm. of lead and used at a focal distance of 75 cm. (Mayneord, 1933; Honeyburne, Lamerton, Smithers and Mayneord, 1939). In particular, in the region of 200 kv. to 400 kv., many measurements have been made, and we may make use of some of these in what follows.

Consider, for example, the results shown in figure 2. These show the distribution of radiation in a plane containing the central axis of a circular beam of x rays sent perpendicularly into a mass of water. The curves are drawn joining places having the same dose, and are, therefore, known as *isodose curves* or *dose contours*, all numbers representing the dose at a given point expressed as a percentage of the dose recorded at the centre of the beam on the surface. This quantity is usually known as *percentage depth dose*. Many hundreds of these sets of curves have been drawn from experimental observations and are of the

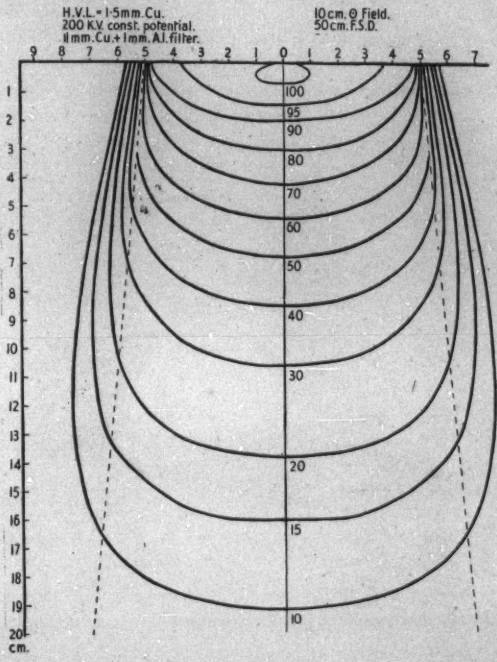


Figure 2. Distribution of radiation throughout a plane containing the central axis of a circular beam of radiation.

utmost practical value in that, as transparencies, they may be laid on cross-sectional diagrams of the body and the dose received at a depth may be estimated for any conditions of irradiation or arrangement of converging beams which are proposed. Modern radiation therapy is increasingly dependent on such distribution diagrams.

Our task is, however, only just beginning, for it is found that the problem as set for the hospital physicist is usually a three-dimensional one of great complexity in which beams of radiation are sent in at all kinds of angles to each other, while the axes of the beams are by no means invariably confined to one

plane. In order to solve the three-dimensional problems, the most promising method of attack seems to be to develop first of all methods of finding the distribution of radiation in any plane however inclined to an incident beam. If the beam is circular, then clearly, on rotation of any distribution diagram such as figure 2 about its axis, each of the isodose curves generates an isodose surface. The determination of the distribution of radiation in any plane, however inclined to the axis of the beam, is now seen to be the determination of the intersections of these isodose surfaces by that plane. A simple geometrical device (figure 3) may be used to determine the required intersections (Honeyburne *et al.*, 1939). The diagram of the radiation of interest may be drawn on glass or Perspex so as to be capable of rotation in a yoke CDC′, while at the same time the axis of the beam AB may be set at any inclination to, and at any height above, a drawing board G on which the required diagram is to be traced. A vertical steel pointer J, ending below in a sharp pencil K, is adjusted to the height of the plane in which the distribution is desired. If the pointer J be

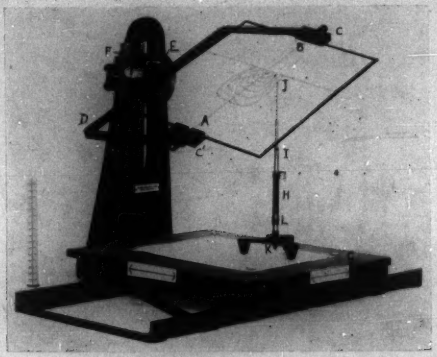


Figure 3.

brought into contact with an isodose curve of the diagram, then on depressing the pencil, a mark is made at a point vertically below a point in the intersection of the isodose surface and the plane in which the pointer moves. By simultaneous rotation of the diagram and adjustment of the pointer of the marker to a given isodose curve the shape of the intersection of the plane with this isodose surface is rapidly traced. In this way the distribution of radiation in any plane may be quickly found, and by studying a number of parallel planes and drawing the diagrams in coloured ink on glass very instructive transparent models of the isodose surfaces may be constructed (Honeyburne *et al.*, 1939).

We have found this contour projector a useful instrument for such rather academic exercises, but for solving the practical problems set by our medical colleagues it is still inadequate. A more powerful weapon is the dose-finder, which is a development of the contour projector.

The dose-finder (figure 4) in its present form (Honeyburne *et al.*, 1939; Mayneord and Honeyburne, 1941) consists of three accurate slides A, at right

angles to each other, carrying a number (usually four) of steel pointers of adjustable length. Suppose now the distribution of radiation is required within a patient to be subjected to several beams: we make a plaster mould (B) of the patient and set up this mould on an adjustable stand so that one of the pointers may be adjusted to reach all points of interest within the mould. A model of the end of the applicator (C) defining the beam is then set in relation to the mould exactly as the real beam will be adjusted on to the patient. The mould is then removed and set in the same relation to the second pointer as formerly to the first. Meanwhile a distribution diagram D of the first beam is slipped into the applicator so as to be capable of rotation. A second model E of the applicator of a possible second beam is now similarly adjusted into position on to the mould (as shown in figure 4) and the process repeated. This process is continued until all the beams are correctly set each to its own pointer. Since all the pointers

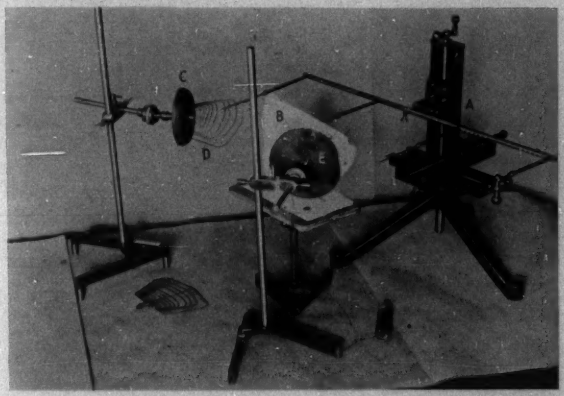


Figure 4. Dose finder. The instrument is shown at a time when a second beam of radiation is being adjusted to a model of the patient.

correspond to the same position in space, to estimate the total dose received we rotate the distribution diagrams until they touch their respective pointers, read off the various doses, and by simple addition (preferably with an electric adding machine) find the total at that point. By movement of the slides the total dose at any position in space may be found, any plane of interest investigated, or any other studies carried out. As an interesting accessory we may use a skeleton or a set of anatomical cross-sections set in relation to yet another pointer and so transform our Cartesian coordinates into anatomical location, to the great relief of our medical colleagues.

The details of the medical applications of these instruments do not concern us here, but we may add that they are equally applicable to the study of radium techniques using teleradium apparatus or the distributions around groups of radium needles, or, indeed, any three-dimensional problems of these kinds (Mayneord and Honeyburne, 1941). The descriptions are merely inserted here

to show how the isodose surfaces may be studied, and in the hope that the instruments may be of interest to physicists faced with similar problems in other fields.

By these methods we may make an estimate of the dose at any point throughout a medium traversed by these radiations, and we now turn to another aspect of the problem, namely, that of the energy absorbed both at a given point and also throughout the medium as a whole. We must therefore now consider the relationship between the röntgen and energy absorption.

§ 5. THE RELATIONSHIP OF THE RÖNTGEN TO ENERGY ABSORPTION

It will be recognized that the röntgen is essentially a measure of energy absorption in air, the factor of proportionality depending upon the energy required to produce one ion pair. This quantity has been extensively investigated (Gray, 1936), and it seems that its value is approximately 33 electron-volts. It follows that a dose of one röntgen corresponds to the absorption of 0·11 erg per 0·001293 gramme of air, or approximately 85 ergs per gramme of air per röntgen. The energy absorption per röntgen in other media may be deduced from theoretical considerations first set out by W. H. Bragg in 1912, who showed that the energy absorption in a medium traversed by penetrating radiation depends upon the real energy absorption coefficient of the medium and the range of the secondary electrons in it. Recent studies by Gray (1937) have placed the theory upon a more exact basis. We may therefore calculate the energy absorption per electron in a number of media of biological interest (Mayneord, 1938), whence it appears that for wave-lengths below 0·1 A., materials of low atomic number are practically identical, but in the region of 0·1 A. to 0·3 A. rapid variation may occur. Beyond 0·3 A., the relative values depend principally on the fluorescent absorption coefficients, and are, therefore, approximately independent of wave-length. It seems that " soft tissues " and " air " are similar to each other, and, therefore, the energy conversion in air is an approximate measure of the energy absorption in soft tissues provided there is electronic equilibrium.

§ 6. RELATIONSHIP BETWEEN INTENSITY AND DOSAGE-RATE

Consider a parallel beam of radiation incident upon an element of mass dm of air. Then if I is the intensity of the beam in ergs/cm.^{2} sec., the energy converted per second into electronic motion is given by

$$I(\sigma_a+\tau)dm\,.\,N, \qquad \ldots\ldots(1)$$

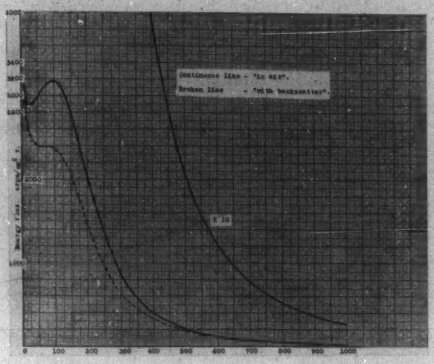
where N is the number of electrons per unit mass of material and σ_a and τ are respectively the real absorption coefficients per electron associated with recoil and photoelectrons. We have seen, however, that a dose of one röntgen delivered to a mass dm of air corresponds to energy absorption 85 dm ergs, so that if D_s is the dosage-rate in röntgens per second,

$$D_s = I\frac{(\sigma_a+\tau)\,.\,N}{85}\ \text{r./sec.} \qquad \ldots\ldots(2)$$

Alternatively, if total energy E_0 flows per unit area in a time t, then the total dose recorded (D) is given by

$$D = E_0(\sigma_a + \tau)\frac{N}{85} \text{ röntgens.} \qquad \ldots\ldots(3)$$

It is of considerable interest to calculate the total flow of energy per square centimetre per röntgen from this equation (Mayneord, 1940a), the result being shown in figure 5. Some of the values given are similar to those of Lauritsen (1933), but cover a wider range of wave-lengths and are displayed rather differently.



λ *in* X.U.

Figure 5. Energy flux per röntgen at different wave-lengths.

It will be seen that in the short wave-length region, the total energy flow per röntgen is of the order of 3000 ergs/cm², but falls to very low values (27 at 1 A.) for long wave-lengths. This distinction between dosage-rate and intensity is of great theoretical and practical importance.

§ 7. TOTAL ENERGY ABSORPTION

Consider now the parallel beam of radiation of cross-section A cm² and of intensity I ergs/cm² sec. entering a mass of absorbing material similar to air for a time t. Imagine, too, that we make distribution measurements throughout the absorbing mass and find for every element of mass dm the dose D_m in röntgens delivered to it. Let ρ be the electronic density of the medium compared with that of air. Then the energy absorbed throughout the scattering mass may be written

$$\text{Total energy absorbed} = \kappa\rho\int^m D_m \,.\, dm, \qquad \ldots\ldots(4)$$

where $\kappa\rho$ is a constant of the order of 85 ergs/gm. röntgen and the integral is taken throughout the absorbing mass. Clearly, if all the energy were absorbed we would expect

$$AIt = \kappa\rho\int^m D_m \,.\, dm. \qquad \ldots\ldots(5)$$

We see, therefore, that the distribution studies already described contain a clue to total energy absorbed in a scattering mass, and we turn, therefore, to the various methods of finding the integral of equation (4). It is convenient to take the energy absorption when one röntgen is delivered to one gramme of air as a unit. We shall refer to it as one gramme-röntgen. The quantity obtained by integrating the dose throughout a given mass we shall refer to as the *integral dose* throughout that region. For medical purposes it is found that a gramme-röntgen is far too small a unit, since we irradiate thousands of grammes with doses of thousands of röntgens, so that one million gramme-röntgens (a megagramme-röntgen) is found to be of a very convenient order of magnitude. If we have a medium of density unity and atomic number equal to that of air, one megagramme-röntgen is approximately equal to 2·02 gm. calories, a convenient conversion factor when estimating total energy absorbed in heat units.

The method given below is the logical extension of the methods of estimating radiation energy from ionization produced in media surrounding the source by Eve (1914), Stahel (1929) and Gray (1936).

§8. ESTIMATION OF INTEGRAL DOSE

Suppose we have determined the distribution of dose in röntgens in a medium of unit density and approximately air-like, as, for example, a water phantom. Then the integral dose in a ring element (figure 6) of cross-sectional area da is clearly

$$D \,.\, 2\pi r \,.\, da.$$

The integral dose between the surfaces generated by the rotation of the contours D_1 and D_2 about the central axis is

$$\Sigma = 2\pi \int r \,.\, D \,.\, da. \qquad \ldots\ldots(6)$$

If D is the mean dose between the contours D_1 and D_2, then

$$\Sigma \simeq D \,.\, V, \qquad \ldots\ldots(6\,a)$$

where V is the volume between the isodose surfaces.

To find this volume (Mayneord, 1940a) we suppose the diagram drawn on cardboard, sheet lead or other suitable material having a mass ρ_s per unit area. Then the moment of the shaded area between D_1 and D_2 about the axis AB is

$$\int r da \rho_s \,.\, g = -V \,.\, \rho_s g/2\pi. \qquad \ldots\ldots(7)$$

If, therefore, we measure the moment of each section of the contour diagram and multiply by the appropriate conversion factor as well as the mean dose, we obtain an approximation to the total absorption in gm. röntgens between the two surfaces.

To measure the moments we have modified a simple chemical balance (figure 7) by fitting on to it a Perspex platform, the moments being determined by weights in the scale pan. Often comparative measurements are all that are required, but if necessary it is easy, by measurements of the length of the beam of the balance and of ρ_s, to determine the moments and volumes in absolute units.

We have tested the method by measuring the moments of simple shapes which generate on rotation solids whose volumes may be directly calculated.

This method is clearly applicable to all beams of x or gamma rays having an axis of symmetry, and has been used by us (Mayneord, 1940a) to determine the integral doses delivered by a number of commonly used techniques, and has given results of very considerable significance in relation to the serious constitutional disturbances experienced by patients undergoing high-voltage x-ray or gamma-ray treatment.

There are many other possible methods of estimating integral dose, though very little experimental work has yet been carried out on the subject. For

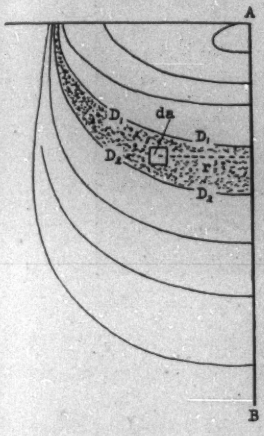


Figure 6.

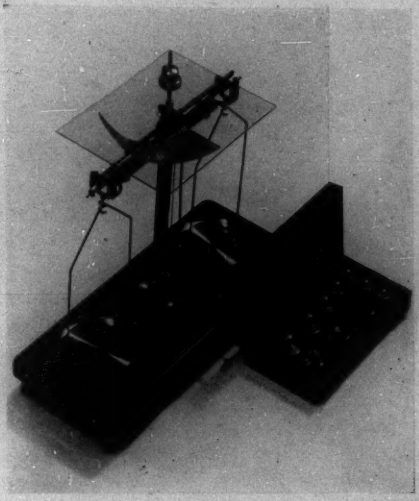


Figure 7.

example, we might send our beams into solutions known to be decomposed by x rays and make use of the chemical changes brought about. Grimmett (1939) has constructed a sectional model of a man specially adapted to measure the total energy absorption throughout his mass, and although no results are so far available, this seems a very direct and promising method of attack.

We have tried to solve the problem in another way by the use of a large number of condenser ionization chambers so placed as to measure the mean dose at the centre of mass of equal masses.* For example, if we suppose the human body to be equivalent in cross-section to an ellipse of semi-axes 18 cm. and 10 cm. respectively, we may divide this ellipse into 72 equal areas, as shown in figure 8. Consider a slice of uniform thickness. The areas then represent

* Condenser chambers suitable for such work, and made by pressing carbon-bakelite mixtures, were demonstrated at the lecture.

equal masses. By placing a condenser chamber at the centre of mass of each section we transform our integral into a summation, so that

$$\int D_m \,.\, dm = \Delta m \,.\, \sum_1^r D_m, \qquad \ldots\ldots(8)$$

and it is merely necessary to add together the indications of all the condensers

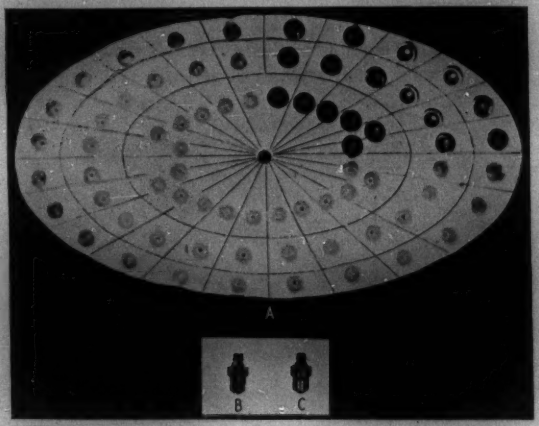


Figure 8. Experimental arrangement for measurement of integral dose.

A. Block of wax divided into 72 equal areas. Ionization chambers in position in one quadrant.
B. Cross-section of condenser chamber.
C. Condenser chamber of pressed bakelite-carbon.

to obtain a quantity proportional to integral dose. By making a number of measurements at different cross-sections we may estimate the total energy absorbed by the body as a whole.

§ 9. SIGNIFICANCE OF INTEGRAL DOSE

It may be wondered why so much interest should be taken in total energy absorption. There are, however, signs that the outlook which has hitherto predominated in dosage problems, namely, that of estimating röntgens at each point within the tissues, has led to an over-simplification of the problem. The energy absorbed by a living cell may not be the sole factor in determining the biological effects of radiation in that cell, for the damage caused to it may also depend on that caused to neighbouring cells which may, in their turn, have been damaged by the radiation. We must remind ourselves that a patient is an organized system, and not a series of disconnected autonomous units.

Moreover, the improvements in precision of the last few years have resulted in raising the dosage levels administered to patients until the severe general constitutional effects of the treatment frequently provide as effective a bar to higher dosage as the local effects previously observed at the skin or other sensitive

tissue. The increase of these general effects with the mass of irradiated tissue, that is, size of field and penetration, is a well-known phenomenon of high-voltage x-ray and radium therapy.

Again, when we consider the physical problems involved in the protection of staff and patients from unwanted radiation, the same difficulties are forced upon us. We might guess that the blood changes and general effects of the radiation would depend, not on the dose at a particular point, but on the total energy absorbed by the body as a whole. This is a very important matter needing much more detailed study, for the Protection Regulations lay down the maximum permissible dosage-rate as 10^{-5} r./sec., whereas we have seen that the total energy flowing through a surface per röntgen recorded on it varies very greatly with the wave-length of the radiation. It is chiefly with the object of studying this problem that the condenser-chamber techniques described above have been devised, for it would seem likely that a given surface dose of gamma rays or penetrating x rays would correspond to much greater total energy absorption in the body than the same dose in röntgens of long-wave diagnostic or therapeutic radiations. Again, the obvious dependence of total energy absorbed in the cross-section of the beam merits more interest than has been given to it. Until further measurements have been made, it is difficult to decide on relative total body absorption per unit surface dose, but it should eventually be possible to find the surface doses of different radiations to give roughly constant total energy absorption.

§ 10. APPROXIMATE FORMULAE FOR THE CALCULATION OF INTEGRAL DOSE

It would be a great convenience to have available formulae for the approximate calculation of integral dose in gramme-röntgens when we do not wish to carry out the elaborate detailed analysis, as, for example, in dealing with the integral doses of hundreds of patients as a routine procedure. Formulae of this nature have been given by Happey (1940, 1941) and the writer (1940 a, 1940 b).

Consider first of all a parallel beam of radiation of cross-section A cm² falling normally on the surface of the absorbing medium. We may to a first approximation consider the dose contours rectangular and reaching to the geometrical limits of the beam. Then, if the fall of dose is exponential (as has been found experimentally to be approximately the case in many instances), and D_0 is the dose on the surface, the dose at a depth x cm. will be given by $D_0 e^{-\mu x}$. Then the integral dose Σ to a depth d in a medium of unit density is given by

$$\Sigma = \int_0^d D_0 . A . e^{-\mu x} dx = \frac{A . D_0}{\mu}\left[1 - e^{-\mu d}\right]. \qquad \ldots\ldots(9)$$

If all the energy is absorbed, $\Sigma = D_0 . A . 1/\mu$. $\qquad \ldots\ldots(10)$

It is interesting that $1/\mu$ is the average path of the quanta in the medium. In practice it is more convenient to use a quantity $d_{\frac{1}{2}}$, the depth at which the dose falls to half its surface value. Clearly

$$\mu d_{\frac{1}{2}} = 0{\cdot}693, \qquad \ldots\ldots(11)$$

$$\Sigma = 1{\cdot}44 . D_0 . A . d_{\frac{1}{2}}, \qquad \ldots\ldots(12)$$

a very convenient formula.

If we consider a diverging beam from a point source at a distance f, the absorption in an element at depth x may be assumed to be

$$\Delta\Sigma = D_0 . A . e^{-\mu x} . dx \left(\frac{f+x}{f}\right)^2. \qquad \ldots\ldots(13)$$

Performing the integration, it is found (Mayneord, 1940 a) that for total absorption

$$\Sigma = 1{\cdot}44\, D_0 . A . d_1^2 \left[1 + \frac{2{\cdot}88\, d_1}{f} + \frac{4{\cdot}16\, d_1^2}{f}\right]. \qquad \ldots\ldots(14)$$

This is the formula we normally employ in practice in calculations on patients, taking into account the dose on the exit surface if appreciable.

We have shown in the same paper that these formulae give results in fair agreement with the more complex detailed study, but much more precise measurement and calculation are now urgently required.

Clearly the same kind of argument may be used for gamma rays, but we are now more concerned with the possibility of correlating the amounts of energy involved with those used in ultra-violet light and infra-red therapeutic procedures.

§ 11. COMPARISON OF ENERGY FLOW OF X RAYS WITH OTHER RADIATIONS

The measurement of ultra-violet light for medical purposes raises very difficult and complex questions which it would certainly be impossible to discuss here in detail, but it may be said that the difficulty arises largely from the very complex variation of biological effect of the same amount of energy with wave-length. This may be exemplified by the variation of erythema-producing power with wave-length according to the investigations of Coblentz, Stair and Hogue (1932). A sharp maximum, set as 100 %, occurs at $\lambda 2976$, coinciding with a strong line of the hot mercury-arc spectrum. We do, however, have some idea of the absolute amount of energy required to produce an erythema on the skin, for this quantity has been estimated a number of times, with, as might be expected, very varying results. Luckiesh, Halladay and Taylor (1930) give 4300 μwatt sec./cm^2 as the mean of their determinations, while Coblentz, Stair and Hogue (1932) give 20,500 μwatt sec./cm^2 as their mean. Maddock (1936) in a paper to this Society found 11,500 μwatt sec./cm^2. The Council of Physical Therapy of the American Medical Association, based largely on the very extensive pioneer work of Coblentz in these matters, defined a ***Finsen Unit*** (F.U.) as a flux density of 10 μwatt/cm^2 of homogeneous radiation of wave-length 2976 A., and say that 2 F.U. for 15 minutes is a representative requirement for a minimum perceptible erythema. This would correspond to 18,000 μwatt sec./cm^2 or $1{\cdot}8 \times 10^5$ ergs/cm^2 at the wave-length of maximum sensitivity. If we use the heterogeneous ultra-violet radiation of a mercury arc (high temperature, high vapour pressure, low voltage, quartz tube) shorter than 3130 A., the total energy required is 58 μwatt/cm^2 for 15 minutes, that is about three times as much energy as for the monochromatic radiation (Coblentz, 1940). It is interesting to calculate the doses in röntgens of x rays of different wave-lengths which would

be required to give the same energy flux as the minimum perceptible erythema dose in ergs/cm^2 at 2967 A. These values are given below:—

λ in X.U.	Dose in *r* without back-scatter
10	50
50	50
150	54
300	175
800	2,800
1,000	5,700
2,000	40,000

It is of interest that in X-ray techniques in which the whole body is irradiated to produce constitutional effects not unlike those of ultra-violet light, the order of dose normally given is 50 to 100 r. at an average wave-length of 150 to 200 X.U., so that the X-ray energy-flux is of the same order as the energy-flux for ultra-violet. This may, of course, be a mere numerical coincidence, but should be further investigated.

These comparisons are also of interest in relation to the cure of rickets, the production of tumours of the skin by both X rays and ultra-violet, and in many other respects.

§ 12. INFRA-RED RADIATION

It would be expected that the amount of energy administered to a patient during treatment by infra-red radiation would be much larger than in either of the radiation fields hitherto mentioned, but very little investigation of the subject seems to have been made, the dosage of infra-red being in clinical practice almost entirely empirical.

We have so far carried out only preliminary measurements, which may now be described.

A Moll microthermopile with a fluorite window was used in conjunction with a Cambridge galvanometer of short period, the e.m.f. generated by the radiation in the thermopile being compensated in a circuit similar to that used by Guild (1930) for radiation measurements. The thermopile was calibrated at the National Physical Laboratory and found to give 61·5 microvolts per milliwatt of radiant energy. A second calibration is also available, as the Bureau of Standards in Washington very kindly presented us with a standard of thermal radiation in the form of a calibrated carbon-filament lamp of a type supplied by them for such work. A careful comparison of the calibrations is in progress.

We have also available small vacuum thermocouples presented by the General Electric Company (Schenectady) in the form of small glass bulbs containing very thin strip thermo-elements (G.E.C., 1941). These are sensitive, quick acting, and when they were used in conjunction with a low-resistance (70-ohm) microammeter, a current of 51 μa. was obtained for a radiation intensity of 1 g. cal./cm^2 min. These vacuum couples were calibrated by a substitution method, using the Moll thermopile as standard. The instruments were mounted side by side on a short optical bench at right angles to the main optical bench

carrying the source of radiation. By moving the instruments along the shorter bench, the General Electric thermoelements could be brought to exactly the position previously occupied by the Moll and so irradiated in similar conditions. The General Electric thermocouple was mounted in a polished metal tube, care being taken to insulate the system thermally so as to reduce conduction to the cold junctions. The thermocouples were calibrated either using the same galvanometer and compensation system or the General Electric instrument connected in series with a low-resistance microammeter, and the currents in this instrument observed directly. The latter procedure is convenient, as the mounted thermocouple and microammeter form a useful clinical instrument with a scale which is very nearly linear up to 100 μa., though there is some departure at higher intensities. Experiments on the wave-length dependence of the calibration are in progress.

Measurements have been made on a number of patients and in conditions likely to occur in practice, the source of radiation during calibration and subsequent use being the same. The results show the urgent need for standardization, since apparently small changes in the source cause large changes in the intensities, while measurements over the surface of the patient or along the axes of the beams reveal very complex variations.

To quote representative results, we have found that, using a 230-volt 500-watt tungsten-filament lamp in its reflector, the amount of radiation received on the skin at a metre distance may vary from 0·58 to 1·8 gm. cal./cm^2 min. The normal value employed for such a source is about 1·5 gm. cal./cm^2 min. We have also found that an intensity of the order of 2 gm. cal./cm^2 min. is about the maximum for comfort, while the threshold of pain is found at about 3·5 gm. cal./cm^2 min., in agreement with Hartley (1934), who gave a value of 3·7 gm. cal./cm^2 min. for a similar source. The value depends, of course, on the part of the body irradiated. The amount which may be tolerated is, however, very variable with the colour temperature of the source, as has been demonstrated by Sonne (1926), who found that the maximum amounts of visible, "inner" and "outer" infra-red which could be borne by human skin were in the ratios of 3·11 to 1·79 to 1·33 gm. cal./cm^2 min., though taking into account reflection, the amounts absorbed were found to be 2·02, 1·16 and 1·33 gm. cal. respectively. These variations may be correlated with the differences in penetration of the radiations. It is interesting to recall that the intensity of sunlight in a clear day in summer at noon is of the order of 1·0 to 1·5 g. cal./cm^2 min.

We hope, in conjunction with our medical colleagues, to continue these measurements on a large number of patients, and so obtain information on the many questions which now arise.

Finally, it may be of interest to see the order of amounts of energy used in the various methods.

(*a*) Type of radiation		Intensity in ergs/cm^2 sec.
High voltage x rays. 200 kv. Heavy filter	40 r./min.	2×10^3
Ultra-violet light	2 Finsen units 20 μ watt/cm^2	2×10^3
Infra-red	1 gm. cal./cm^2 min.	7×10^5

(*b*) Type of radiation		Total energy absorbed
x rays. 200 kv.	30 megagramme-röntgens	60 gm. cal. in six weeks
Ultra-violet light	Erythema over body surface of 1500 cm^2	8 gm. cal. at one sitting
Infra-red	1 gm. cal./cm^2 min. 20 min. Area 1500 cm^2	3×10^4 gm. cal. at one sitting

§ 13. CONCLUSIONS

We have seen how the measurement of radiation for medical purposes is concerned primarily with two energy-absorption problems, first the energy absorbed at each point in the tissues, and secondly the energy absorbed in the body as a whole. Concerning the first problem, in the x-ray and gamma-ray field we now know a little, but so far as ultra-violet and infra-red are concerned, practically nothing.

As regards the total energy absorbed by the body as a whole, we have now rough estimates which are of great interest and enable us to link up many different problems.

Only on an energy basis can these problems be correlated, and it is hoped that this sketch of a very wide field may do something to stimulate interest in the measurement of radiation in absolute units for medical purposes.

REFERENCES

Bragg, W. H., 1912. *Studies in Radioactivity* (Macmillan and Co., Ltd.).

Coblentz, W. W., 1940. *J. Amer. Med. Assoc.* **114**, 325.

Coblentz, W. W., Stair, R. and Hogue, J. M., 1932. *Bur. Stand. J. Res., Wash.*, **8**, 541, 759.

Eisl, A., 1929. *Ann. Phys., Lpz.*, **3**, 227.

Eve, A. S., 1914. *Phil. Mag.* **27**, 394.

G.E.C., 1941. *Gen. Elect. Rev.* **44**, 133.

Gray, L. H., 1936. *Proc. Roy. Soc.* A, **156**, 578.

Gray, L. H., 1937. *Brit. J. Radiol.* **10**, 600; **10**, 721.

Grimmett, L. G., 1939. *Amer. J. Roent.* **41**, 432.

Guild, J., 1930. *J. Sci. Instrum.* **7**, 378.

Happey, F., 1940. *Nature, Lond.*, **145**, 668; **146**, 96; 1941. *Brit. J. Radiol.* **14**, 235.

Hartley, H., 1934. *J. Inst. Fuel*, April.

Honeyburne, J., Lamerton, L. F., Smithers, D. W. and Mayneord, W. V., 1939. *Brit. J. Radiol.* **12**, 262.

Kaye, G. W. C., Bell, G. E., Binks, W. and Perry, W. E., 1939. *Rep. Progr. Phys.* **6**, 95.

Lamerton, L. F., 1940. *Brit. J. Radiol.* **13**, 136.

Lauritsen, C. C., 1933. *Amer. J. Roent.* **30**, 380, 529.

Luckiesh, M., Halladay, L. L. and Taylor, A. H., 1930. *J. Opt. Soc. Amer.* **20**, 423.

Maddock, A. J., 1936. *Proc. Phys. Soc.* **48**, 57.

Mayneord, W. V., 1933. *Brit. J. Radiol.* **6**, 598.

Mayneord, W. V., 1938. *Rep. Progr. Phys.* **5**, 284.

Mayneord, W. V., 1940 a. *Brit. J. Radiol.* **13**, 235; 1940 b. *Nature, Lond.*, **145**, 972.

Mayneord, W. V. and Honeyburne, J., 1941. *Amer. J. Roent.* **45**, 235.

Mayneord, W. V. and Lamerton, L. F., 1941. *Brit. J. Radiol.* **14**, 255.

Mayneord, W. V. and Roberts, J. E., 1937. *Brit. J. Radiol.* **10**, 365.

Sonne, C., 1926. *Acta radiol., Stockh.*, **5**, 419.

Stahel, E., 1929. *Strahlentherapie*, **33**, 296.

Taylor, L. S. and Singer, G., 1940. *Amer. J. Roent.* **44**, 428.

DIELECTRIC LOSS IN PARAFFIN-WAX SOLUTIONS *

By H. FRÖHLICH,
H. H. Wills Physical Laboratory, University of Bristol

MS. received 7 *March* 1942

ABSTRACT. In this paper, the dielectric loss of a dilute solution of long-chain molecules, containing dipoles, in solid paraffin is investigated theoretically. In such substances, realized experimentally by Jackson, Sillars and Pelmore, the interaction between dipoles is negligible, and they thus follow the Debye theory. In order to account for the time of relaxation τ, the structure of these substances is investigated and it is shown that only the case where the dipole-containing chain molecule is shorter than the paraffin molecule can be expected to yield simple results. In order to obtain τ, the minimum energy V required to turn a molecule from one equilibrium position to another has been calculated. For short molecules, V is mainly determined by the potential hill between two equilibrium positions, whereas for long molecules an internal torsion of the molecule becomes important. V is found to increase linearly with chain length for very short molecules, and to tend towards an asymptotic value for long chains. Good agreement with experiments, so far as available at present, has been reached.

§ 1. INTRODUCTION

It is well known experimentally that dielectric losses are very small in solid insulators which form ionic or atomic crystals. Appreciable dielectric losses occur mainly in molecular crystals, built up of molecules containing electrical dipoles.

The fact that dipoles are in many cases responsible for dielectric losses was recognized by Debye (1929) and led to his well-known theory of this effect. Applied to solids, the main idea of this theory is as follows: suppose that each dipole has two positions of equilibrium corresponding to opposite dipole directions with equal energy, and a certain transition probability from one to the other. Now assume that an external electric field is being applied. Then, clearly, some of the dipoles will turn into more favourable directions, thus giving rise to a dielectric polarization. Owing to the finite transition probability, it will take some time to reach equilibrium. This time is measured by the time of relaxation τ. Obviously, if the external field is periodic, the polarization will also be periodic, but with a certain shift of phase. This, according to Maxwell's equations, gives rise to dielectric losses.

Although there seems to be little doubt that the general ideas of Debye's theory are correct, detailed comparisons with the results of experiments on solids have met with varying success. This is probably due to the fact that the interaction between dipoles has not been described in a satisfactory way.

* Based on Report L/T. 121 of the British Electrical and Allied Industries Research Association (E.R.A.).

In order to measure the time of relaxation τ of chain molecules without meeting the difficulties just mentioned, W. Jackson (1935) diluted polar long-chain molecules in paraffin wax. Pure paraffin wax, being non-polar, shows practically no dielectric loss. The losses in the solid solution, therefore, can be ascribed to the dissolved polar molecules. Moreover, if their concentration is small enough, their interaction will be negligible.

In this way, Jackson (1935), and later R. W. Sillars (1938) and D. R. Pelmore (1939), determined the time of relaxation of long-chain esters of various chain lengths. Their results show that with increasing chain length, $\log \tau$ increases too, but slower than linearly, and probably approaches an upper limit for very long chains. Now the significance of $\log \tau$ is as follows: Let V_n be the minimum energy which is required in order to lift a molecule with n links over the potential hill between the two stable positions. Then, since $1/\tau$ represents the transition probability, it may be expected that

$$\tau = Ce^{V_n/kT}, \qquad \ldots\ldots(1)$$

where, compared with $e^{V_n/kT}$, C varies only slowly with temperature and chain length. Assuming the molecule to be rigid, V_n, and hence $\log \tau$, must be proportional to the chain length. The experiments, therefore, show that this is not the case, but that for long chains V_n approaches a constant value, independent of the chain length.

It is the aim of this paper to calculate the dependence of V_n on the chain length. Allowing for internal rotational deformations of the molecule, it will be shown that this dependence has the simple form of a hyperbolic tangent. The result, which is demonstrated in figure 3, shows good agreement with experiment. It ought to be mentioned that this result holds only if the dipolar molecule is shorter than the paraffin molecule. In § 2 this will be derived from a study of the lattice structure of paraffins. In § 3, V_n will be calculated, and the results will be discussed in § 4.

§ 2. STRUCTURE

We shall start with a discussion of the possible positions which the long-chain ester molecules may be expected to take under the assumption that they form a real solid solution in the paraffin crystal. The structure of paraffins is well known from the work of A. Müller (1928, 1932). To describe it, we divide the crystal into plane layers, each of the thickness of the chain length. Within a layer, the unit cell is rectangular, with sides of lengths b_1, b_2 and c. b_1, which is of the order of 7·5 A., and b_2 (5 A.) are parallel to the surface of the layer; c is perpendicular to it, and its length is equal to that of the chain. All chains are parallel to the c-axis. Figure 1 shows the way in which they cross the b_1b_2-axis. It is very important to notice that, in the subsequent layer, the whole arrangement is shifted by about 1 A. in the b_1 direction. These atoms are indicated by dotted lines in figure 1. Now let us try to build a foreign chain into the paraffin crystal. This chain must obviously be parallel to the paraffin chains. It may either replace a paraffin chain, or it may be in an interstitial position, probably in the middle of the b_1-axis. It does not seem possible to decide at present which of the two alternatives is realized, but for the following considerations this is

unimportant. Now clearly we have to distinguish two cases, one where the length of the foreign molecule is shorter and the other where it is longer than the paraffin chain. In the first case, the foreign molecule can be built in between the paraffin chains of one layer without being deformed. In the second case this is no longer possible, because of the above-mentioned displacement of two successive layers relative to each other. Such a long molecule has, therefore, to be deformed, in order to be built into the crystal. There is some experimental

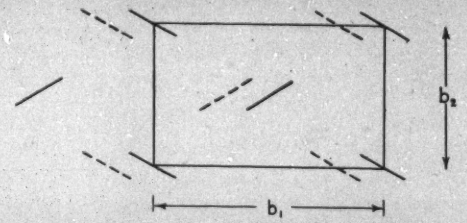


Figure 1. Structure of paraffins perpendicular to chain axis.

evidence for the difference between the two cases. X-ray measurements by Piper and others (1931, 1937) show that the crystal spacing is hardly changed if we add to a crystal of chain molecules a small concentration of shorter molecules, whereas, in contrast, the addition of longer molecules increases the spacing considerably.

Now let us consider the equilibrium position of a chain in more detail. Any such chain forms a plane zigzag, and, as indicated in figure 1, these zigzag

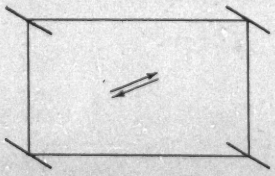


Figure 2. The two equilibrium positions of a dipolar chain.

planes are in certain fixed directions.* The same must hold for the zigzag plane of a foreign molecule. If such a molecule contains a dipole, then there are at least two equilibrium positions with equal energy but different dipole directions. This is shown in figure 2 under the assumption that the foreign molecule replaces a paraffin chain. The arrows indicate the projection of the dipole direction on the $b_1 b_2$ plane.

§ 3. CALCULATION OF V_n

Let us try to turn the dipolar molecule from one equilibrium position to the other, by rotating it at one end. Then, assuming it to be rigid, we have at the same time to lift the whole molecule to the top of the potential hill which

* At high temperatures, slightly below the melting point, many paraffins change into a modification in which the chains probably rotate. We shall restrict ourselves, however, to the more important low-temperature modifications.

separates the two stable positions. If, however, we allow for an internal elasticity, the molecule will be twisted. In order to simplify the following calculations, we shall treat the chain as a continuum, described by elastic constants.

Let z be the distance of any point of the chain axis from one end, and suppose that a CH_2 group with the co-ordinate z has been turned from one of its equilibrium positions by an angle ϕ around the z-axis. Assuming the chain to be a continuum, ϕ is a function $\phi(z)$ of z, and we shall have to calculate the energy of a deformation in terms of ϕ. Thus, the potential energy will be $\propto \phi^2$, owing to the interaction with its neighbouring molecules. If ϕ depends on z, the chain is twisted, and there will be an additional internal elastic energy $\propto \left(\frac{d\phi}{dz}\right)^2$. Thus the total energy is given by

$$\int_0^{l_n} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\}\frac{dz}{a}. \qquad \ldots\ldots(2)$$

Here l_n is the chain length, and a is the distance between two links, i.e.

$$an = l_n; \qquad \ldots\ldots(3)$$

A^2 and B^2 have the dimensions of an energy, whose order of magnitude we shall estimate below. According to figure 2, with $\phi \equiv 0$, $\phi \equiv \pi$ is also an equilibrium position of the chain. The energy of a small deformation from this latter position is obtained from (2) by replacing ϕ by $\phi - \pi$.

In order to be able to carry out quantitative calculations, we shall assume the ϕ^2 law to hold in the interval $-\frac{1}{2}\pi \leqslant \phi \leqslant \frac{1}{2}\pi$ and the $(\phi-\pi)^2$ law in $\frac{1}{2}\pi \leqslant \phi \leqslant \frac{1}{2}\pi$. Then the interaction energy has its maximum at $\pi/2$ and at $-\pi/2(=3\pi/2)$. Now the most general position of the molecule having an energy near V_n must be such that $\phi = \pi/2$ at a certain $z = z_0$ with both ends $z=0$ and $z=l_n$ hanging down towards $\phi=0$ and $\phi=\pi$. The potential energy W of the molecule in such a position is given by

$$W = \int_0^{z_0} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\}\frac{dz}{a} + \int_{z_0}^{l_n} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}(\phi-\pi)^2 \right\}\frac{dz}{a}. \qquad \ldots\ldots(4)$$

V_n is to be determined from the condition that W is stationary, i.e.,

$$\delta W = 0. \qquad \ldots\ldots(5)$$

The variation of W comprises a variation of the function $\phi(z)$ as well as of z_0. ϕ has to be determined so as to make W a minimum for any given z_0. Then, starting to turn the chain at $z_0 = l_n$, i.e. lifting the one end of the chain from $\phi=0$ up to $\phi=\pi/2$, we shall move z_0 towards $z=0$, which means that we gradually move the chain over the potential hill, always using the most suitable $\phi(z)$. It follows at once from symmetry that the energy thus determined reaches its maximum at $z_0 = l_n/2$. Hence V_n is obtained from (5) with this value of z_0. Using this value of z_0, we obtain from (5) and (4)

$$a^2A^2\frac{d^2\phi}{dz^2} - B^2\phi = 0, \text{ if } 0 \leqslant z \leqslant \frac{l_n}{2}, \qquad \ldots\ldots(6a)$$

and

$$a^2A^2\frac{d^2\phi}{dz^2} - B^2(\phi-\pi) = 0, \text{ if } \frac{l_n}{2} \leqslant z \leqslant l_n, \qquad \ldots\ldots(6b)$$

with the condition that

$$\phi = \pi/2 \text{ at } z = l_n/2, \qquad \ldots\ldots(7)$$

and the condition for free ends, viz.,

$$\frac{d\phi}{dz} = 0 \text{ at } z = 0, \text{ and } z = l_n. \qquad \ldots\ldots(8)$$

We can now calculate V_n, which is given by

$$V_n = \int_0^{l_n/2} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\} \frac{dz}{a} + \int_{l_n/2}^{l_n} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\} \frac{dz}{a},$$

where ϕ satisfies (6), (7) and (8). Using these equations, we find by a partial integration

$$V_n = A^2a\left(\phi\frac{d\phi}{dz}\right)_{z=l_n/2} = A^2a\frac{\pi}{2}\left(\frac{d\phi}{dz}\right)_{z=l_n/2}. \qquad \ldots\ldots(9)$$

To determine $d\phi/dz$ we have to solve the differential equations (6). Using the boundary conditions (7) and (8) we obtain

$$\phi = \frac{\pi}{2}\frac{e^{z/a\bar{n}} + e^{-z/a\bar{n}}}{e^{n/2\bar{n}} + e^{-n/2\bar{n}}}, \text{ if } 0 \leqslant z \leqslant l_n/2,$$

and

$$\pi - \phi = \frac{\pi}{2}\frac{e^{(z-an)/a\bar{n}} + e^{-(z-an)/a\bar{n}}}{e^{n/2\bar{n}} + e^{-n/2\bar{n}}}, \text{ if } \frac{l_n}{2} \leqslant z \leqslant l_n,$$

where

$$\bar{n} = A/B. \qquad \ldots\ldots(10)$$

Therefore

$$\left(\frac{d\phi}{dz}\right)_{z=l_n/2} = \frac{\pi}{2a\bar{n}}\frac{e^{n/2\bar{n}} - e^{-n/2\bar{n}}}{e^{n/2\bar{n}} + e^{-n/2\bar{n}}}.$$

Inserting this value into (9), we finally obtain

$$V_n = \frac{\pi^2}{4}AB \tanh\frac{n}{2\bar{n}} = \frac{\pi^2}{4}B^2\bar{n}\tanh\frac{n}{2\bar{n}}. \qquad \ldots\ldots(11)$$

We thus see that, for short chains, $n \ll \bar{n}$, V_n becomes

$$V_n \simeq \frac{B^2n}{2}\frac{\pi^2}{4}, \quad n \ll n,$$

as would be obtained if the molecule were considered to be rigid. For long chains, $n \gg \bar{n}$, however, V_n becomes

$$V_n \simeq AB\,\pi^2/4, \quad n \gg \bar{n},$$

independent of the chain length. In this case, therefore, the torsion of the chain is of importance.

§4. DISCUSSION

If we insert equation (11) into (1) we find for the time of relaxation τ

$$\log 2\pi\tau = \alpha + \frac{n^2}{4}\frac{B^2\bar{n}}{kT}\tanh\frac{n}{2\bar{n}}. \qquad \ldots\ldots(12)$$

This formula contains three constants: α, B and $\bar{n} = A/B$. We shall not attempt here to determine α theoretically, but the order of magnitude of B and $\bar{n}$ can be estimated as follows: from its definition, B^2 is the energy required per link

in order to turn the untwisted chain by an angle $\phi = 2$ ($\sim\pi/2$). This should be about the same energy as is required per link in order to turn a paraffin chain in the same way. Since at about 300° K. paraffins have a transition into a state in which their molecules can probably rotate, we estimate

$$B^2 \sim 1/30 \text{ ev.} \qquad \ldots\ldots(13)$$

A^2 determines the energy required to twist the chain, i.e. to deform the valency angle. Thus A^2 may be expected to be of the same order of magnitude as the energy of the chemical bond. Thus we expect

$$A^2 \sim 5 \text{ ev.} \qquad \ldots\ldots(14)$$

Hence according to (10)

$$\bar{n} = A/B \sim 10. \qquad \ldots\ldots(15)$$

Before discussing our results in a more quantitative way we must recall that in § 2 we assumed that the dipolar molecule is shorter than the paraffin chain. If it is longer it will have to be strongly deformed in order to be built into the lattice. No simple theoretical significance can be given to such measurements. With these remarks in mind, we shall now compare quantitatively the experimental and theoretical dependence of $\log\tau$ on the chain length.

From Sillars' experiments we know $\log\tau$ for $n =$ 20, 22, 24, 28 and 32 in ordinary paraffin wax. Since the chain length of the latter is between 24 and 28, only the values for $n = 20$, 22 and possibly 24 can be used. Measurements by Pelmore (1939) are, however, available for $n = 32$ in very long paraffin chains (up to about 52 links).

In the following table we give the experimental values of $-\log_{10}(2\pi\tau)$ at $-20°$ C. The values in brackets are those which have to be excluded:

n	20	22	24	28	32	
$-\log_{10}(2\pi\tau)$	+6·7	+5·5	4·8	(4·2)	(4)	(Sillars)
					2·3	(Pelmore)

In order to determine the unknown constants in the theoretical formula (12) we shall use the experiments on the temperature dependence of τ. Unfortunately, for the substances for which our theory holds, the measured temperature range is too small to check the theoretical temperature dependence. Assuming it to be correct, it is possible to determine (from $n = 20$) two constants, namely,

$$\alpha = -50{\cdot}4, \quad \frac{B^2\bar{n}}{k} = 5200°. \qquad \ldots\ldots(16)$$

The remaining constant $\bar{n}$ can be determined from the ratio of τ for $n = 20$ and for $n = 22$, which gives

$$\bar{n} = 13. \qquad \ldots\ldots(17)$$

Thus from (16) and (10) the values of A^2 and B^2 are

$$A^2 \sim 6 \text{ ev.}, \quad B^2 \sim 1/30 \text{ ev.}, \qquad \ldots\ldots(18)$$

in agreement with the theoretically estimated order of magnitude (14) and (13).

In figure 3 we have plotted the theoretical curve, i.e. $\tanh n/2\bar{n}$, as a function of $n/2\bar{n}$. Using the above values for α, A^2 and B^2, we see that the experimental

points for $n=24$ and $n=32$ lie on the theoretical curve, whereas the values for $n = 28$ and 32, which had to be excluded (in brackets), fall below.

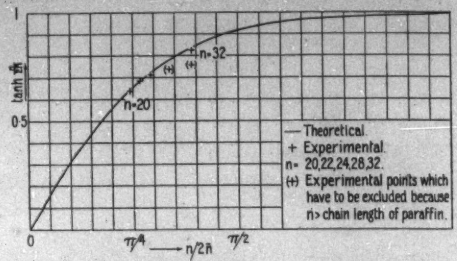


Figure 3. Dependence of $\log \tau$ on chain length.

The agreement reached can be considered as a confirmation of our formula (12). It is satisfactory to notice that the value of $B^2\bar{n}$ obtained from the temperature dependence agrees with that obtained from the dependence on chain length, and has, moreover, the theoretically estimated order of magnitude.

REFERENCES

DEBYE, 1929. *Polare Molekeln* (Leipzig).
FRANCIS, COLLINS and PIPER, 1937. *Proc. Roy. Soc.* A, **158**, 691.
JACKSON, W., 1935. *Proc. Roy. Soc.* A, **150**, 197.
MÜLLER, A., 1928. *Proc. Roy. Soc.* A, **120**, 437 ; 1932. *Ibid.* **130**, 514.
PELMORE, D. R., 1939. *Proc. Roy. Soc.* A, **172**, 502.
PIPER, CHIBNALL, HOPKINS, SMITH and WILLIAMS, 1931. *Biochem. J.* **25**, 2072.
SILLARS, R. W., 1938. *Proc. Roy. Soc.* A, **169**, 66.

THE DIMENSIONS OF PHYSICAL QUANTITIES

BY E. W. H. SELWYN

Communication H849 from the Kodak Research Laboratories, Harrow

Received 20 *March* 1942

ABSTRACT. It is shown that the simple form of relation between physical quantities which is assumed in considerations of the dimensions of physical quantities may be regarded as a consequence of the equality of two measures of a physical quantity, on any system of units, if they are equal on one system. The argument is extended to cover the " method of dimensions ".

§ 1. INTRODUCTION

A FEW years ago, in a theoretical discussion of the origin of the graininess of photographic materials (Selwyn, 1935), an argument was developed which involved the supposition that photographic density could be regarded as having a dimension in a way similar to that in which mass or time has a dimension. Now photographic density is defined as the logarithm of the ratio of two light fluxes or even of two transmissions. Thus, in the ordinary

way, it would undoubtedly be regarded as of zero dimensions. But the assumption that it had a dimension (namely, the dimension of density) did, in fact, lead to an answer in agreement with answers obtained by straightforward physical arguments, and, indeed, in substantial agreement with observation. Moreover, the form of the expression arrived at, by whatever process, forces upon one the conclusion that the " graininess constant " which forms a convenient measure of the granularity of photographic materials is of the dimensions [*photographic density*] . [*length*]. It was felt, by analogy with certain other physical " constants ", such as specific heat, or the quantity a in the equation

$$T = 2\pi\sqrt{l/g}(1 + a\theta^2)$$

for the period of a pendulum vibrating through a small finite arc θ, which has different values accordingly as the amplitude is measured in radians or degrees, that there was justification enough for this assumption that density had a dimension of its own. The problem was therefore worked out by the ordinary " method of dimensions ". Then Bridgman's *Dimensional Analysis* (Bridgman, 1931) came to the author's notice, and the arguments there put forward were immediately adapted to suit the special case in question, since they appeared much more convincing. The treatment which now follows of the general question of dimensions is broadly that followed by Bridgman,* but it is believed to have sufficient points of novelty, particularly in the derivation of the fundamental formulae, to justify publication.

§ 2. THE FUNDAMENTAL FORMULAE

In measuring a physical quantity, the process adopted is to measure certain other quantities, say, $a, b, c, d, \ldots$, and combine them by a specified method of calculation to yield the desired result, say p. Thus

$$p = F(a, b, c, d, \ldots). \qquad \ldots\ldots(1)$$

If the quantity being measured is a simple one, such as a length for instance, the process of measurement and calculation is very simple. One may regard the process of measuring a length as consisting in taking a rod and finding out how many times the rod can be placed end to end within the length which is being measured. The number of times so counted is called the length, while the rod with which the measurement is made is called the unit of length. Obviously, if the unit were shortened to a fraction $1/L$ of its original length, the number which represents the measured length would be multiplied by L. If we suppose that all the quantities in the brackets of (1) are such that this reciprocal relation exists between the measures and the units, and we change the units in the ratios $1/A$ for the a-quantity, $1/B$ for the b-quantity, and so on, we shall find a new value for the p-quantity, namely,

$$p' = F(Aa, Bb, Cc, Dd \ldots). \qquad \ldots\ldots(2)$$

It is customary to repeat physical measurements several times, and we now suppose that a second measurement of p has been made, with different values $a', b', c', d', \ldots$ as the starting point, but with the result remaining the same

* Resemblances will also be found to that given by Campbell (1920).

as before, namely p. We now impose the very reasonable condition that the result of the calculation after the units of the a'-, b'-, c'-, d'-, ... quantities have been changed in the ratios $1/A$, $1/B$, ... shall remain p'. That is, we make the requirement that if two measurements of a physical quantity are found to be equal on one system of units, they shall remain equal on any other system. It would be difficult to find a less restrictive condition. In other words, we expect that

$$\left.\begin{array}{ll} \text{if} & F(a, b, c, d, \ldots .) \text{ is constant,} \\ \text{so also is} & F(Aa, Bb, Cc, Dd, \ldots .), \end{array}\right\} \quad \ldots\ldots(3)$$

whatever the values of A, B, C, D. This constitutes a limitation and, as will be seen immediately, a very considerable limitation on the form of F, that is upon the types of calculation permissible in deriving physical quantities.

Let us consider the simplest case, in which

$$p = F(a, b) \quad \ldots\ldots(4)$$

$$\text{and} \quad p' = F(Aa, Bb). \quad \ldots\ldots(5)$$

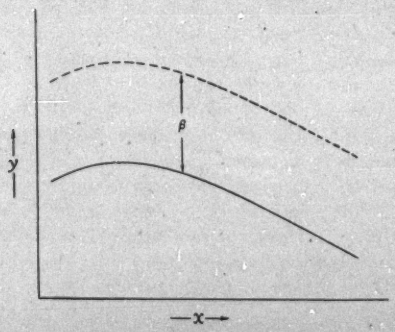
First change the variables to $x = \log a$, $y = \log b$, $\alpha = \log A$, $\beta = \log B$ and $p = z$, $p' = z'$. Then we can write the above equations as

$$z = f(x, y),$$
$$z' = f(x + \alpha, y + \beta).$$

We consider these expressions to represent surfaces whose height above the plane of the x-y axes is z, and in the figure we take the full line to represent the contour line

$$f(x, y) = z_0.$$

Now for every point (x, y) on this curve, draw a corresponding point $(x, y + \beta)$. The curve through the points $(x, y + \beta)$ will obviously be parallel to the curve through the points (x, y). But according to the preceding arguments, if $f(x, y)$



is constant, so also is $f(x, y + \beta)$. That is, since $f(x, y)$ or $f(x, y + \beta)$ is the height of the surface above the plane of the x-y axes, the second curve is also a contour line. Thus, by these arguments, either z is independent of x and y (a trivial solution), or all contour lines are parallel. The same arguments can be applied

to changes of a-units, and also to simultaneous changes of both a- and b-units. These conditions are clearly only satisfied if the contour lines are straight, for with any parallel lines which are not straight, a displacement can be found which will cause them to intersect. Thus

$$z = f(lx + my), \qquad \ldots\ldots(6)$$

where l and m are constants, and f represents an arbitrary function.

An exactly similar argument can be used in discussing the " contour-surfaces " involved in the four-dimensional case, for which the result is

$$z = f(lx + my + nz). \qquad \ldots\ldots(7)$$

This result can be extended to any number of co-ordinates by taking them in groups of any three, while maintaining the remainder constant. Since all groups of any three must follow the preceding equation, the final result after changing back to the original variables is that

$$p = F(a^\alpha b^\beta c^\gamma d^\delta \ldots.). \qquad \ldots\ldots(8)$$

Bridgman suggests that it is required of physical quantities that they shall preserve the same ratio when the units of measurement are changed. We can arrange for this by taking the inverse function, F^{-1}, of p as the measure of the quantity in question, for then its measure becomes

$$a^\alpha b^\beta c^\gamma d^\delta \ldots. \qquad \ldots\ldots(9)$$

and in different units

$$A^\alpha B^\beta C^\gamma D^\delta \ldots. a^\alpha b^\beta c^\gamma d^\delta \ldots \qquad \ldots\ldots(10)$$

The convenience of this convention is obvious. There seems to be no very fundamental reason, however, why we should not find it convenient, on occasion, to use quantities defined by

$$F(a^\alpha b^\beta c^\gamma d^\delta \ldots.).$$

In fact, it is common to do so. For some purposes, for instance, the logarithm of a brightness is a better unit than brightness itself. But it must be remembered that if brightness measurements are to be combined with other measurements in deriving a physical quantity, it will be found simpler to use brightness itself rather than its logarithm. The a-, b-, c-, d-quantities involved in the derivation of a p-quantity are therefore themselves invariably reduced to such a form that they can be expressed by formulae of the type of (9) before they are used in the calculation. Moreover, so long as the form of F is known, it is immaterial whether we calculate the quantity we are interested in as $a^\alpha b^\beta c^\gamma d^\delta \ldots.$ or as $F(a^\alpha b^\beta c\ d^\delta \ldots.)$. From this point, therefore, we shall assume the simpler form.

The above demonstration implicitly assumes that a, b, c, d, ... are all independent. But it may happen that the units of the individual components of a group of quantities, say $e^\varepsilon f^\zeta g^\eta$, are such that $E^\varepsilon F^\zeta G^\eta$ is independent of the values of E, F, G. This occurs when e, f, g have component measures in common, the indices of which in the product $e^\varepsilon f^\zeta g^\eta$ all add to zero. A product of this sort is termed dimensionless, because a change in the size of the units

involved in it is without effect on its value. Consequently, as a little consideration will show, in these circumstances,

$$p = a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta} \ldots . \phi\, (e^{\varepsilon}\, f^{\zeta}\, g^{\eta}),$$

and in general

$$p = a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta} \ldots . \phi\, (a', b', c', d' \ldots .),$$

where the arguments a', b', c', d', ... of ϕ are each dimensionless products similar to $e^{\varepsilon}\, f^{\zeta}\, g^{\eta}$. These products must, of course, be independent.

Some p-quantities are constant. Typical examples are

$$G = d^2 F / m_1 m_2,$$

where d is the distance between two masses m_1, m_2 and F the force of attraction between them; and

$$a = u/\Theta^4,$$

where u is the energy density of radiation in a black-body enclosure at absolute temperature Θ. Constants such as these have different values according to the units used in the measurement of the component factors, and are commonly known as dimensional constants. Other constants, such as

$$\pi = C/D,$$

where C is the circumference and D the diameter of circular objects, and " Reynolds number ",

$$v \rho d / \eta,$$

where v is the velocity of a fluid, η the viscosity, ρ the density of the fluid, and d a linear dimension describing the size of the channel in which the fluid is flowing, are independent of the units used in its measurement, and may be conveniently termed " function constants ". Now these constants may be regarded as substitutes for certain measurements and must, therefore, be included, in certain fairly obvious circumstances, in the formula for p. To take a very simple example, we might be interested to know the relation between the diameter of a wheel and the distance travelled by it in so many turns. The formula for the calculation involves π, because the quantity we measure is D, but the quantity which determines the distance travelled is C. In the same way, if we wish to know the volume of a liquid drop which falls from a tube, among other things we know that one of the important factors is the force on the drop tending to pull it out of the tube. We may measure this in our experiments with the drop itself by finding how fast it falls and the weight of the drop, or by making use of the constant g for gravitational acceleration. There is, however, a fundamental difference between the two sorts of constant. A " function constant " is necessarily included among the arguments of ϕ, but a dimensional constant can be involved in ϕ only if it is combined with such other terms as to make the resultant dimensionless.

We have now reached the general result that the complete formula for a physical quantity is

$$A^{\alpha}\, B^{\beta}\, C^{\gamma}\, D^{\delta} \ldots . a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta} \ldots . \phi\, (a', b', c', d' \ldots .),$$

wherein $A^{\alpha}\, B^{\beta}\, C^{\gamma}\, D^{\delta} \ldots .$ represents what are known as the " dimensions " of the quantity, but where the rest of the formula represents what " physical

reality" the quantity has. The insignificance of $A^\alpha B^\beta C^\gamma D^\delta \ldots$ is emphasized by the fact that it is hardly ever left in physical equations. There are two reasons why it is left out. One is that there are no absolute units by which $a, b, c, d \ldots$ might be measured. One set of units is therefore just as appropriately used for $a, b, c, d \ldots$ as another. The other is that it can always be determined without the slightest difficulty from the normal formula

$$p = a^\alpha b^\beta c^\gamma d^\delta \ldots \phi(a', b', c', d' \ldots).$$

§ 3. THE METHOD OF DIMENSIONS

A reverse operation is possible to some extent. This constitutes the basis of the "method of dimensions". The fundamental assumption underlying this process of arriving at relations between physical quantities is that the dimensional formula for the p-quantity may be known independently of the way in which the relevant factors determine it in the experiment actually under consideration. This is, indeed, a necessary condition for most of the operations of physics. For the experimental determination of the relation between p and its component measurements $a, b, c, \ldots$ presupposes a knowledge of p by other means. We might, for instance, define the surface tension of a liquid by means of measurements of the volume of drops falling from a tube, but this would mean nothing very much unless the same results (within reasonable limits) were also secured for surface tension by measurements on ripples on the surface of the liquid or the height of ascent of the liquid in capillary tubes. Thus we may start off with a knowledge of the dimensions of p in terms of certain units a_0, b_0, c_0, d_0, and the suspicion that p is also measurable in terms of the units $a, b, c, d, \ldots$. We know that if this is so

$$p = a^\alpha b^\beta c^\gamma \ldots \phi(a, b, c, \ldots),$$

and we also know the measures of $a, b, c \ldots$ in terms of $a_0, b_0, c_0 \ldots$, namely,

$$a = a_0^{\alpha_1} b_0^{\beta_1} c_0^{\gamma_1} \ldots$$
$$b = a_0^{\alpha_2} b_0^{\beta_2} c_0^{\gamma_2} \ldots$$
etc.

Thus,

$$p = (a_0^{\alpha_1} b_0^{\beta_1} c_0^{\gamma_1} \ldots)^\alpha (a_0^{\alpha_2} b_0^{\beta_2} c_0^{\gamma_2} \ldots)^\beta \ldots \phi(a, b, c, \ldots),$$

and the dimensional part of the formula for p is

$$(A_0^{\alpha_1} B_0^{\beta_1} C_0^{\gamma_1} \ldots)^\alpha (A_0^{\alpha_2} B_0^{\beta_2} C_0^{\gamma_2} \ldots)^\beta \ldots$$

The process of the "method of dimensions" is then to make this formula identical with the known dimensional formula for p. This is done by making the exponents of $A_0, B_0, C_0 \ldots$ the same in the above formula as in the already known dimensional formula.

Sometimes this is not possible. Then we know that there can be no relation whatever between the variables we have chosen and p. If just one single set of exponents $\alpha, \beta, \gamma, \ldots$ makes the formulae agree, then we conclude that if a relation such as we have assumed does exist, it necessarily has the form we have found. But this by no means implies that the relation does exist. For instance, one might assume that the height H of a human being who had grown up on the

moon would depend upon the gravitational acceleration, g, there, and the time he had taken to grow, t. Then

$$H = \text{const. } gt^2,$$

as may easily be found by the usual procedure. But this is undoubtedly a wrong answer. If H is the distance a stone drops, starting from rest, in a time t, however, it is a correct answer physically. Frequently, however, there is more than one set of exponents which will make both dimensional formulae equal. Then an arbitrary function of arguments of zero dimensions finds its way into the answer. An example may, perhaps, be better than a formal demonstration. Let us consider how the distance H travelled by a falling body varies with the time t if the acceleration due to gravity is g and the starting velocity is v. We can, in this case, choose the ordinary mechanical units of mass, length and time for our a_0, b_0, c_0 . . . . units. Then we assume

$$H = g^\alpha v^\beta t^\gamma \phi$$

with the dimensional relation

$$[L] = [LT^{-2}]^\alpha . [LT^{-1}]^\beta [T]^\gamma.$$

Now we can make a dimensionless product out of the variables g, v, t, namely, v/gt. Consequently we know that our formula will finally be

$$H = g^\alpha v^\beta t^\gamma \phi(v/gt).$$

Considering only the dimensional part for the moment, however, we find by equating coefficients that

$$\alpha + \beta = 1,$$
$$2\alpha + \beta - \gamma = 0.$$

Consequently, we have arrived only at a relation between the coefficients and not a unique solution. But the above equations show that

$$\alpha = 1 - \beta,$$
$$\gamma = 2 - \beta,$$

whence

$$[L] = [LT^{-2}] . [T^2] \{[LT^{-1}]/[LT^{-2}] . [T]\}^\beta,$$

which shows a striking similarity to *

$$H = gt^2\phi(v/gt).$$

The occurrence of the indeterminable coefficient β is due to the possibility of forming dimensionless products of the variables, and in working out the relations between the indices we infallibly find these dimensionless products. This may readily be seen from the consideration that if a dimensionless product exists, such as

$$[LT^{-1}]/[LT^{-2}] . [T],$$

the power to which it is raised has no effect whatever on the equality of both sides of the equation—in respect of dimensions—and the exponent is therefore

* The normal expression $H = gt^2/2 + vt$ is recognizable as in agreement with this expression if it is written $H = gt^2(1/2 + v/gt)$.

not determinable. But the connection between this indeterminacy and the arbitrary ϕ, which we have already shown to be part of the physical equation, appears to be only incidental. It is probably worth remarking again that the relation obtained by the method of dimensions is not necessarily a sensible one. How sensible the relation is will, in fact, depend upon how sensible we have been in considering and formulating the problem.

§ 4. GENERAL CONCLUSIONS

All that we have required of a physical quantity for it to be expressible by means of the simple formula always adopted in considerations of dimensions is that if two measures of it are equal on one system of units they shall also be equal on any other system. So long as the components which appear in the formula are measured on scales which preserve the same ratio between two measures of a component whatever the units of the scale, any physical quantity can be regarded as having dimensions. There is no restriction whatever to mass, length and time dimensions. In using the " method of dimensions ", for instance, we are at liberty to use any system of units which can be consistently applied to the definition of the quantity we are measuring and the quantities in terms of which we are trying to express it. The intrinsic reality of such quantities or any such esoteric considerations are of no moment.

The only question at issue is : are they necessary in the calculation of the quantity we are interested in? It is completely immaterial, for instance, that the Newtonian constant of gravitation and the inverse-square law are only an approximation; the use of the Newtonian constant in dimensional arguments, about planetary motion for instance, certainly leads to the right answer.

REFERENCES

BRIDGMAN, P. W., 1931. *Dimensional Analysis* (2nd Edn.) (Yale Univ. Press).
CAMPBELL, N. R., 1920. *Physics : The Elements*, Ch. 13 and 14 (Cambridge Univ. Press).
SELWYN, E. W. H., 1935. *Phot. J.* **75**, 571-80.

AN ELECTROMETER FOR MEASUREMENT OF VOLTAGE ON SMALL IONIZATION CHAMBERS

BY F. T. FARMER, PH.D.,
Barnato Joel Laboratories, The Middlesex Hospital

Communicated by Prof. S. Russ, 27 March 1942

SMALL condenser ionization chambers, such as the Sievert type, are used extensively in radiological and other work for the measurement of gamma-ray intensities. These have a very small capacity, often as low as 1 $\mu\mu$F. or less, and the measurement of voltage drop due to exposure necessitates the use of a low-capacity electrometer. A fibre instrument such as the Wulf or Lindemann is usually employed for this purpose.

The expense of such an instrument, and the fact that a somewhat elaborate optical system is needed if an external scale reading is to be provided, suggested

that some simpler system was worth investigating, particularly for routine measuring purposes. An electrometer valve may be used as the basis of such an arrangement, and it has been found that the device described below, which gives a direct reading on the dial of a meter as a measure of chamber voltage, meets the requirements. The accuracy is as high as that of the electrostatic instruments, and the more compact and robust nature of the apparatus is an advantage also in certain cases, such as where measurements in a hospital radium department are required.

Essentially, the highly insulated grid of the valve takes the place of the electrometer suspension fibre, so that, if the internal electrode of the chamber is connected directly to this while the outer electrode is at a fixed potential, the current through the valve provides a measure of the voltage on the chamber.

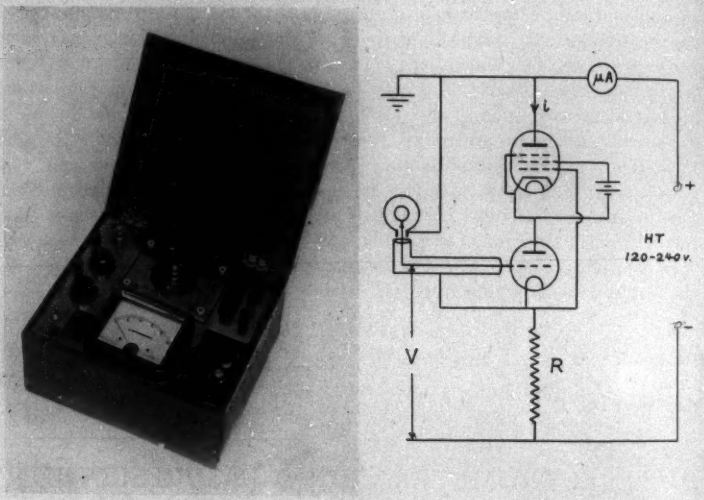


Figure 1.

The voltages which are to be measured are, however, of the order of 100 to 200 volts, so that means must clearly be provided either for reducing the effective voltage on the chamber to that which can be embraced by the grid base of the valve, or for extending by some method the working range of the valve itself. The former could be readily achieved by connecting a fixed capacity in parallel with the chamber, so that the charge of the latter is shared with this, and the potential so reduced to any desired value. The second method, however, has much more important possibilities, since it admits the use of a large degree of negative feed-back, and this can be put to advantage to render the whole measuring system relatively independent of supply voltages as well as of variations in the valve characteristics. The considerable precautions which would be necessary, if the former method were used, to ensure constancy of voltage

to anode and filament may thus be done away with, and, as will be seen from the calibration curve (figure 2), even a change from 120 to 240 volts in the high-tension battery does not alter the instrument reading by more than 1 %. Similar considerations apply to the filament supply, and an arrangement which requires the minimum of adjustment for precise measurement is thus rendered practicable.

The electrometer valve must be operated with an anode voltage not exceeding 4 or 5 volts, and, therefore, it cannot be employed simply with a cathode resistance to provide feed-back, on account of the large voltages which would appear at its anode. To meet this difficulty, a second valve of high impedance is placed in series with it, as shown in figure 1. The bias potential for the grid of this valve is obtained directly from the voltage drop in the electrometer valve itself, and in consequence the potential difference between anode and filament of the latter is kept to its required low value. The encumbrance of the second valve is not great, since for the small current which it has to pass, only a few volts are required to supply its screen, and its filament may be operated from the same source as the electrometer.*

The two valves together function in the manner of a " cathode follower ", and it is clear that the relationship between grid potential, V, relative to the negative terminal of the battery, and the reading, i, of the microammeter, is given very nearly by

$$i = V/R,$$

where R is the cathode resistance. The characteristics of the valves and batteries only enter to a second order, and approximate linearity of the overall calibration comes about also through these considerations. A cathode resistance of the order of 1 megohm, allowing a maximum current of about 100 micro-amperes to flow through the instrument, is employed.

The inherent capacity of the grid and its connecting lead to the chamber needs to be taken into account, since the charge to be measured is shared with this. With the valve used, a Marconi E.T.1, it amounted to 3 $\mu\mu$F., which is about three times the chamber capacity, and if no steps were taken to counteract it, it would mean that only a quarter of the voltage on the chamber would appear at the grid, and the negative feed-back could not be made as effective as is desired. To overcome this difficulty, the valve and grid lead were surrounded by a screen connected to the electrometer filament, the potential of which follows very closely that of the grid itself; the capacity to earth, therefore, of the grid and lead is reduced to a very small fraction of its normal value, and almost the whole voltage on the chamber appears at the grid as is desired.

In using the instrument, the chamber must be charged negatively relatively to its case, so that when it is applied to the electrometer, the sudden change in the grid potential is in a negative, not a positive, direction. This is necessary, since, in the transient condition, the cathode potential of the valve lags very slightly behind that of the grid, owing to the appreciable capacity to earth of the filament battery, and if the grid is allowed to assume a more positive potential than its equilibrium value, grid current flows, and the charge is incorrectly

* A 2-volt valve (type 15) with indirectly heated cathode allows this simplification with battery-operated filaments.

registered. A momentary change to a more negative potential, however, causes no flow of current, and, therefore, no disturbance to the working.

In order that the reading of the meter should remain steady after contact with the chamber is made, it is essential that the grid current in the static working condition should be extremely small. The valve mentioned above has been found very satisfactory from this point of view, no perceptible drift of the meter occurring in the time required for taking readings.

Changes in the supply voltages to the instrument are, as pointed out, immaterial unless they are severe, so that no provision is necessary for adjusting high-tension or low-tension voltages. The small current taken from the high-tension battery means, moreover, that there is little tendency for it to run down

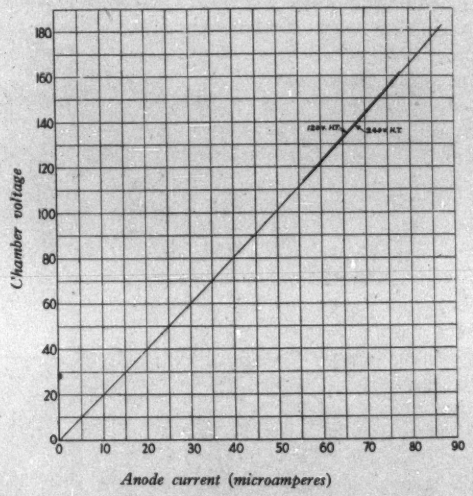


Figure 2. Voltage calibration for Sievert chamber, type (1–1).

except at long intervals. A magnetically operated switch is incorporated for "earthing" the grid of the electrometer valve before it is brought in contact with the chamber, so that it should start from the same potential in each measurement.

A feature of this type of apparatus worth mentioning is that it is not harmed by applying greatly excessive charges to the electrometer valve grid, as may occur when the chambers are used in experimental work. The possibility of damage to a sensitive fibre electrometer through such causes is a factor which has constantly to be borne in mind when such instruments are used.

I am very greatly indebted to Professor S. Russ for providing all facilities necessary for this work; and also to Miss D. Clephan for the photograph reproduced herewith.

A HIGH-VOLTAGE APPARATUS FOR ATOMIC DISINTEGRATION EXPERIMENTS

By J. D. CRAGGS, Ph.D., M.Sc., A.Inst.P.,
Metropolitan-Vickers Electrical Co., Ltd., Manchester

MS. received 2 *February* 1942 ; *in revised form* 13 *April* 1942

ABSTRACT. A simple apparatus for performing experiments in nuclear physics has been developed to the stage where beam currents of 100 to 200 μa. at 600 kv. have been obtained. Experiments with the acceleration tube have led to some interesting results ; the presence of large currents of secondary electrons has been confirmed and their effect on the running of the apparatus is discussed in some detail. It was hoped to carry out experiments on the installation of different forms of limiting diaphragm in the tube, but lack of time and the wish to obtain some results with proton bombardment prevented our doing so.

Some experiments with nuclear γ-rays are briefly described. The results are not accurate, but they have indicated the practicability of continuing such work.

The performance of the tube compares very favourably with other installations using electrostatic generators for the supply of high-tension current. The output from the ion source is capable of being raised almost by a factor of 10, so that if a more powerful generator had been available, target currents of the order 500 μa. should have been easily obtained with a heat dissipation higher probably by a factor 4 than that encountered in the above work. This presents only a minor technical problem.

§1. INTRODUCTION

Since the pioneer work of Cockcroft and collaborators, Tuve, and others, there have been many investigations in the field of artificial atomic disintegrations, and nuclear structure has now become a specialized branch of physics, with its own established techniques. One of the principal requirements in the production of artificial nuclear reactions is an apparatus for accelerating charged particles (usually protons or deuterons) to high speeds, and in large numbers. Disintegrations can, of course, be produced with fairly slow ions ($\sim$50 kv.), but the reaction efficiencies are low, as penetration of the nuclear barriers is a probability effect which increases rapidly with increasing speed of the bombarding particles. Hence, until voltages comparable with the barrier heights are attained, the apparatus should work at as high a voltage as possible. Further, since the disintegration output will vary linearly with beam current, the effect of a low current may be offset by a smaller relative increase in voltage which, in general, is more easily obtained. The limit, of course, is set by the size of building and the increasing cost of the apparatus.

In designing the present apparatus in outline, emphasis was also laid on the need for flexibility—for instance, the use of a Wilson chamber at the target end of the tube instead of Geiger counter or ionization chamber detectors. It is also important to facilitate the irradiation of samples with the greatest possible efficiency, i.e., to use a large target solid-angle, as many induced reactions are of low efficiency—for example, production of photo-nuclear effects by γ-ray bombardment.

§2. REQUIREMENTS FOR THE PRESENT APPARATUS

It was decided to build a neutron generator on modest lines with an output voltage of 700 kv. and a beam current of 200 to 500 μa., as this seemed an economical proposition. Available space was limited by the use of a room

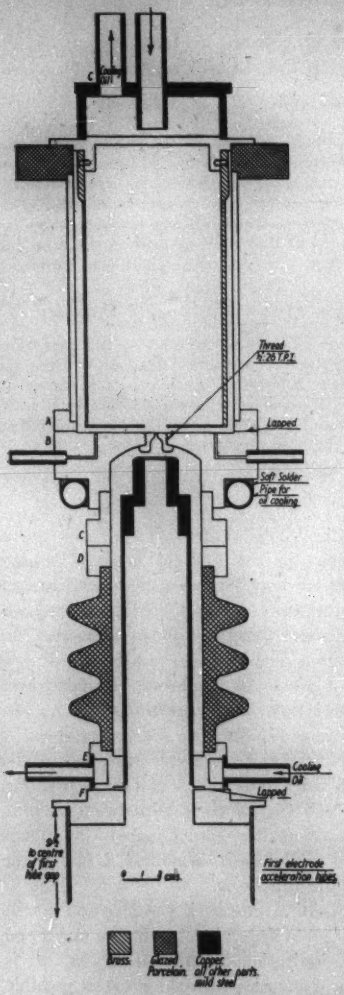


Figure 1. The ion source and focusing electrode assembly. To avoid complicating the diagram the following details are omitted: (1) Bakelite clamping bolts and nuts for ion source with brackets on B and C; (2) centring screws bearing on A with brackets on B; (3) centring screws bearing on C with brackets on D; (4) bakelite clamping bolts with brackets on B and F; (5) centring screws bearing on E with brackets on F; (6) steel clamping bolts with brackets on F and first plate of acceleration tube. Interior surfaces of ion source are polished.

approximately 16 ft. wide × 23 ft. high in King's College, London. The sideways clearance was the controlling factor. During the development period in Manchester, therefore, these figures were carefully considered. Owing to

the present circumstances, the apparatus has not been installed in London, and the preliminary programme of work has been completed in the Metropolitan-Vickers Works at Trafford Park, Manchester.

The apparatus could have been designed with the target at a high, not earth, potential. Before discussing the reasons for adopting the latter arrangement, a brief description of the essential components of the equipment will be given. It is necessary, first, to produce an intense beam of positive ions—a cold-cathode discharge tube was used in our case (Oliphant and Rutherford, 1933)—and then to focus the beam into the main accelerating tube, after which the diffusion and scattering losses should be small, so that the ions proceed to the target. It is possible to use the first one or two stages of the cascade tube for focusing, but the interposition of a separate focusing stage is desirable. The electrode assembly at the inlet end of the acceleration tube is shown in figure 1. The source and focusing will, it is hoped, be described in separate publications. If the source were arranged to be at or near earth potential, manipulation of the controls would be simplified. Further, the pumping equipment, which is necessarily at earth potential, would be utilized to the best advantage, because the gas flowing from the source through the canal into the acceleration tube would be removed near the canal. However, the target would be at a potential of 700 kv. with respect to earth, and many measurements, e.g. counting experiments or the use of a Wilson chamber, would be complicated. On the other hand, if the target were earthed and the source at high potential, as in our arrangement, the pumps would be somewhat choked by the acceleration tube, and the power supplies for source and focusing would have to be remotely controlled. The latter scheme, on account of the large high-tension structures, is more expensive but, on the whole, more convenient, because the target end of the tube is of more importance in running than the source, which, once adjusted, rarely needs attention. The earthed-source arrangement is seldom encountered.

§ 3. DETAILS OF CONSTRUCTION

The following is a brief description of the method of construction of the high-tension electrodes. Wooden frames (2 in. × 2 in. members) were bolted to channel-iron base structures (3 in. × 2 in. × $\frac{3}{8}$ in.), which in turn were secured to the bakelite cylinders. Finally, the structures were covered with $\frac{1}{32}$-in. sheet steel, the end caps being mounted first, with steel supporting bars, 1 in. × $\frac{1}{8}$ in., attached to the wooden framework. The final appearance of the apparatus is seen in figure 2. The van de Graaff generator is 112 in. high, its terminal being 57 in. in diameter. The height of the electrodes is about 4 ft. 6 in. The bakelite cylinders are 75 in. high, 7$\frac{3}{4}$ in. internal diameter and $\frac{3}{4}$ in. thick. In figure 2 metre scales are shown, fixed to the right-hand front support of the large electrode and to the front ground girder.

The design of the frames was carefully considered. The total weight carried by the two sets of bakelite supports was about 1$\frac{1}{2}$ tons. The bakelite cylinders were adequately strong: the framework of the larger electrode was the most highly stressed part. Even so, it was estimated that the safety factor was about 8, which was considered adequate. Vibrations are only appreciable when the driving motor, on slowing down after a run, forces the electrode into resonance, but the effect is not serious.

It was necessary to determine satisfactory values for the electrode radii in order to minimize corona losses, as distinct from increasing the spark-over voltages. The former were important because of the limited current output of the van de Graaff generator. From Peek's (1929) data for cylinder/cylinder gaps it is possible to deduce the maximum gradients obtaining in cylinder/plane gaps: the latter are relevant to a study of the practical case of our high-voltage electrodes. Taking a total voltage of 850 kv. for a gap of 7 ft., and cylinders of 6 in. radius, the maximum gradient is approximately 29 kv./cm., which is near the breakdown value for normal air, viz. 30 kv./cm. As we did not expect to exceed 700 kv. in our earlier work, 8-in. radii were considered sufficient.

The minimum clearance to the laboratory walls (see figure 2) was about 6 ft. to a projecting girder: although this spacing is approximately the same as that for the electrode/ground, the latter is of less importance, because the

Figure 2.

floor provides a large conducting surface which acts as a screen. The driving motor and belt seem to exert no appreciable influence. The electrode coverings (figure 2) were secured to a hoop system of steel strips with screws, the heads of which projected outwards. These have never brushed appreciably, because of the screening effect of the large metal surfaces.

Figure 3 shows corona losses for the complete apparatus. The beneficial effect of thin pressboard insulating screens interposed between the wall and the high-voltage electrode is to be noted. This technique has been used in many laboratories (e.g. by Bouwers at Eindhoven and at Cambridge). The effects of introducing such barriers into asymmetric air and oil gaps were described many years ago (Goodlet, Edwards and Perry, 1931). The breakdown voltages of gaps which do not form corona before spark-over are in general decreased by the interposition of a barrier of higher S.I.C. than the initial medium. However, in the presence of even small corona currents, the reverse holds, i.e. spark-over

voltages are increased, due, presumably, to the barrier becoming charged to a potential near, and of the same sign as, the high-potential electrode, and so acting as a screen. The tentative theory is supported by the fact that the efficacy of the barrier is somewhat increased as it is made to approach the high-potential electrode. There seems to be an optimum screen position for most gaps (Goodlet, Edwards and Perry, *loc. cit.*). The total available output from our van de Graaff generator was about 1·5 to 1·8 ma. positive excitation, at about 700 kv.

The arrangement of equipment in the electrodes is shown diagrammatically in figure 4. The lead to the van de Graaff generator terminal is on the extreme right-hand side. (See also figure 2.) In the larger electrode there are two D.C. sets (+50 kv. and −75 kv. maximum) each "earthed" to the frame, which is at the potential of the ion source (discharge-tube) cathode. The focusing electrode (see figure 1) and the first cylinder in the cascade acceleration tube are connected to the small electrode. The arrangement whereby the

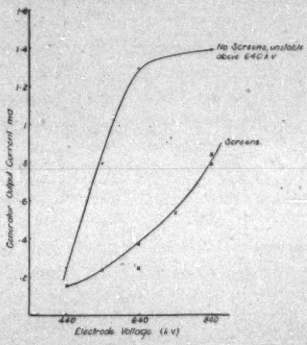


Figure 3. The effect of insulating screens on corona lines from high-voltage apparatus.

cathode is at the potential of the common terminal of the D.C. sets was adopted for technical reasons, e.g. the avoidance of D.C. sets giving voltages >75 kv., which were considered high enough for focusing the ion beam, as the discharge tube was intended to operate at ~30 kv. with a possible 50 kv. maximum. The D.C. sets are remotely controlled by bakelite rods, chain-and-sprocket driven at the control end, and secured to the auto-transformers controlling the high-tension transformer primaries in the electrode. On-off contactors (fitted with overload relays, fuses, etc.) are operated by strings. The ion-source anode is cooled by special oil of low viscosity flowing in insulating tubes through the upper busbar (figure 4) and impelled by a small centrifugal pump driven off the 9·5 kva. alternator in the electrode, which is driven by a 12·5-H.P. motor at earth potential (shown in figure 2): a cooling system, fan and radiator, for the oil is similarly driven. The source and focusing voltages are read from milli-ammeters, measuring the currents flowing through special high resistances in parallel with the electrodes (see § 6), mounted on a central panel with signal lights, and visible through an aperture cut in the electrode. The small electrode

contains, when required, a magnet for separation of unwanted ions from the beam. Power is supplied by a D.C. generator driven by a 2·5-H.P. motor and controlled rheostatically. Usually the magnet was not required, and was then removed to simplify control. Alternatively, the magnet may be mounted at the earthed end of the tube.

The ripple on the focusing and source voltages was $\gtrless 5$ %. The only point of interest in the D.C. sets was the horizontal rectifier valve mounting, adopted to reduce the electrode height to a minimum, and incorporating rubber pads to damp out vibrations.

The electrode system and mounting of the cascade tube are normal (see § 7). However, in order to economize in head-room, the tube was inclined to the vertical, a procedure which increases the difficulty of supporting the tube, adjusting it, and making the joints. After traversing the acceleration tube, the ion beam passes through the tee-piece connected to the diffusion pumps, a $1\frac{1}{4}$-in.

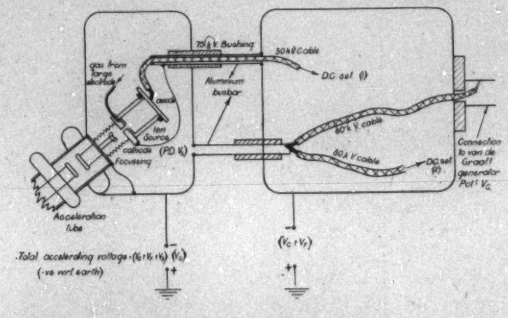


Figure 4. Potential distribution in neutron-generator electrode.

V_F = focusing P.D. $\sim$ 10/70 kv.
V_S = source P.D. $\sim$ 10/20 kv.
V_G = van de Graaff potential ($\sim$ 600 kv.).

bore conical tap, and finally strikes the target, attached to a copper holder. The latter is of conventional design, cut away at 45° to the axis and water-cooled. The techniques adopted for using different targets are described below. Figure 2 shows the corona shields on the tube, the diameters of which were decided upon after calculations similar to those described in some detail for the large electrode.

The van de Graaff generator is controlled from a separate cubicle, and the voltage applied to the large electrode is measured with a generating voltmeter of simplified design. The latter is calibrated with a sphere gap checked in the usual way by observations on gamma-ray resonances, which were supplemented by tests carried out with a set of wire-wound high resistances: we estimate that the error of voltage measurement is about ± 2 %. This was considered satisfactory for preliminary work, and the need for extra stabilization has never arisen.

§4. PUMPING SYSTEM

Because of the appreciable quantity of gas, about 300 c.c. of hydrogen per second at 0·1 to 0·2 mm. Hg pressure, flowing from the ion source, high-speed pumps must be used to give low pressures in the acceleration tube. The mean free path for any likely collision processes should be at least equal to the tube length. Rough extrapolations for ionization and *Umladung* losses in hydrogen (Gerthsen, 1930; Meyer, 1937) indicate mean free paths, at 10^{-4} mm. Hg pressure, for 200-kv. protons to be ~1 m. and ~30 m. respectively. The tube length was about 80 in. (2 m.), to give suitable external gap lengths for 700 kv. operating voltage. Hence the tube pressure should be ~10^{-4} mm. Hg, and preferably 10^{-5} mm. Hg. This value agrees reasonably with published data (Bouwers *et al.*, 1937; Lauritsen *et al.*, 1941), and from the data given above necessitates a pumping speed of 200 to 1000 l./sec.

The pumping set, as finally developed, consisted of a 6-in. diameter brass pump, based on a design by Crane, in series with a 3-in. pump, a Metropolitan-Vickers type 02 and a Cenco-Megavac rotary pump. The diffusion pumps were mounted in an angle-iron frame adjustable for height with corner bolts. The main intake, connected to the acceleration tube, is 6 in. in diameter and is bolted down to the 6-in. pump, using a lead or rubber gasket. The pump speed at the entrance to the acceleration tube was about 300 l./sec. for hydrogen. Subsequent experience has shown this performance to be adequate.

§5. DETAILS OF MISCELLANEOUS ACCESSORIES

Before describing the experiments carried out with the complete apparatus, it was decided to include a section giving details of some of the component parts of the apparatus, where they might be considered useful to other experimenters.

The resistance voltmeters used for measurement of source and focusing voltages are of interest (see §4 for mention of their installation, etc.). An accuracy of ± 5% was considered sufficient. Further, it was required to use robust instruments which could be easily read at distances of about 10 ft.

Specially constructed wire-wound resistances of considerable accuracy, up to about 150 megohms in value, have been used regularly in this laboratory for many years, but were considered unnecessarily exact for this work. We used chains of Morganite 1 or 2 watt resistances arranged in a helix round bakelite supports, and immersed in transformer oil to give adequate cooling and freedom from corona losses.

Galvanometers or electrostatic voltmeters were rejected as indicators in favour of 1-ma. meters. The usual technique adopted in these laboratories is to run such resistances at much less than their rated dissipation because of their high negative temperature coefficients, but we decided, size being a very important consideration, to use currents of the same order as the nominal ratings. For the 50-kv. source voltmeter, 50 1-megohm, 1-watt (1-ma.) resistances were used, as it was realized that 30 kv. would be rarely exceeded: for the 75-kv. focus stage voltmeter, the resistance chain consisted of 15 1-megohm, 2-watt (1·4-ma.) resistances. The spiral assemblies, with brass end plates, fitted closely inside 8-in. diameter bakelite containers and were 9½ in. and 15 in. long respectively.

The nominal 50-megohm unit (50·5 megohms at 20° c.) changed by <3 % at 0·5 ma. in the steady condition. At 0·75 ma., when stability was reached in about 5 hours, the resistance change was about -6 % for a temperature rise of 19 to 20° c. The resistance was linear and free from appreciable hysteresis effects. It was shown by special experiments that the resistance changes were due to the Joule heating and not to any voltage effect. The effect of oil immersion on resistance was shown to be negligible. Further experiments performed after a lapse of two months showed the resistance change to be <1 %.

A test with the nominal 75-megohm resistance (actually 78·6 megohms at 20° c.) gave -10 % change for 1 ma. current, which was considered excessive. There are several standard methods for reducing such effects, e.g. thermostatic control or the use of a potentiometer arrangement whereby the voltage measured is a fraction (independent of temperature changes if the heating of all the resistance units is equal) of the total. A method was devised which seems to be simpler than those just described: it consists of immersing a metallic resistance R_2 in the oil container, so that R_2 reached approximately the same temperature as the main resistance R_1, and applying a shunt R_3 of negligible temperature coefficient across the meter and R_2 in series. R_2 increases as R_1 decreases, so giving a compensation effect. The meter resistance should be much less than R_2 to avoid swamping the compensation effect, and R_2 should equal R_3 for greatest sensitivity. Typical values were:—meter resistance $=40$ ohms, $R_3 \simeq R_2 \simeq 1000$ ohms. Without attempting to obtain the best conditions, i.e. equality of positive and negative temperature coefficients, the effective change in R_1 was reduced to the acceptable value of -4 % for 30° c. rise in temperature, i.e. at 1 ma. current, using a copper coil for R_2.

The measurement of the van de Graaff terminal voltage presents some points of interest. The method used should involve as small a current consumption as possible, and should give continuous readings with stability and ~ 2 % accuracy. The last requirements eliminate sphere gaps, and a generating voltmeter (Kirkpatrick, 1932) becomes the obvious choice. The operating principle is, briefly, as follows. An insulated plate is alternately exposed to and shielded from the high-voltage electrode whose potential we require to measure when shielded, e.g. by a rotating vane system, the pick-up plate loses its charge through a resistance. If the latter be connected to a suitable galvanometer or valve amplifier, a direct reading of voltage is obtained. The instrument is not absolute, although it becomes so in carefully defined geometrical conditions, and is usually calibrated with a spark gap and with γ-ray resonances where the voltmeter is used with atomic disintegration apparatus. A simple single-stage valve amplifier was used (10 to 20 megohm grid leak, depending on required sensitivity) giving 0 to 500 kv. linearly above a small initial exponential part of the curve, with a 0 to 1 ma. meter. A $\frac{1}{8}$-H.P. synchronous motor (1500 r.p.m.) was used to drive the rotor.

It is not possible for the induced changes on the pick-up plate to be measured directly with a D.C. instrument, because the direction of the current changes when the pick-up plate is covered, i.e. twice every revolution. If C is the capacity of the H.T. electrode/pick-up plate assembly with voltage V applied to the electrode, then if Q be the charge at any instant,

$$\frac{dQ}{dt} = i = V\frac{dC}{dt}.$$

dC/dt changes sign when the pick-up plate is completely covered, so that the mean value of dC/dt is zero.

The R.M.S. voltage (E) across the load resistance (R) is shown (van Atta *et al.*, 1936) to be

$$E = 3{\cdot}93 \times 10^{-10}\ F.A.S.R.\ \text{millivolts},$$

where S = rotor speed (r.p.s.), A = effective rotor area, and F = field strength at the rotor in volts/cm. In one case $S = 25$, $A \simeq 160$, $F \simeq 4 \times 10^3$, $R = 10^7$, in the appropriate units. Hence $E \simeq 60$ volts. This is higher than the value obtained in practice by a factor ~ 10, probably because F is much lower than the value taken here, which equals the average gradient from the generating voltmeter to the high-potential electrode. It is clear that the method is capable of great sensitivity. Over an operation period of about six months the voltmeter has shown great stability.

As an even simpler alternative to the above, we attempted to make high liquid resistances ($\sim 10^{10}\,\omega$) using special liquid mixtures in Pyrex containers. The best resistances showed steady changes of about 1 %/day, due probably to differential evaporation of the rather complex fluid. Such resistances would have been satisfactory if re-filled and calibrated once a week, but cannot be recommended for precision work. In view of the interest attached to this problem a separate publication has been prepared and is now in the press.

§ 6. EXPERIMENTS ON THE ACCELERATION TUBE

The tube construction may be seen in figure 2. The cylinders ($^1/_{16}$ in. thick brass) are $9\frac{3}{8}$ in. long, 4 in. internal diameter, and are provided with steps to prevent sliding through the steel end plates which project beyond the glass (6 in. internal diameter, $^3/_{16}$ in. thickness, and 10 in. length) cylinders. The brass electrodes are maintained in alignment because of their driving fit in the end plates, the bearing surface being $\frac{3}{8}$ in. long. The stress distributors mounted on the ends of the electrodes, inside the accelerator tube, were formed from copper tube of $\frac{3}{4}$ in. outside diameter, arranged to give an inter-electrode spacing of $\frac{1}{2}$ in. This last setting is not critical except, possibly, in the first stage of the tube, as the focusing is obtained largely by the use of the primary focusing electrode (figure 1). The glass cylinders are clamped between the end plates by the use of $\frac{1}{2}$-in. diameter Tufnol rods and nuts. All vacuum joints on the tube were made of Apiezon Q compound. The external stress distributors (figure 2) were graded from 7 in. to $1\frac{1}{8}$ in. diameter, the latter at the earthed (target) end of the tube. External spark-over for a single gap with $1\frac{1}{8}$-in. diameter shields was 140 to 150 kv. The alignment of the tube was excellent, as the glass cylinders had been carefully ground to close limits. Lapped flanges with centring screws are provided at the inlet and exit ends of the tube. The ion beam passed through a 12·2 mm.-diameter collimating aperture before passing through the tap to the target.

The first experiments showed the existence of large currents of secondary electrons produced by the relatively small number of stray positive ions in the

tube. This phenomenon had been expected, on the basis of other work with high-speed positive ions (e.g. Hill *et al.*, 1939). For instance, at 240 kv. (negligible corona) some experiments showed a total tube current of 0·6 ma. when the target current was 20 μa. and the entrant positive-ion current $\sim$200 μa.

Clearly this problem is of great importance where the largest possible target current is required with a limited available output current from the van de Graaff generator. Some time was spent in investigating the tube behaviour, as the published data seemed negligible despite the inevitable appearance of the effect in most of the tubes described (e.g. Gray *et al.*, 1940). The most obvious remedy is to prevent the secondary electrons from leaving the regions in which they are produced. In a well focused tube, most of the secondaries are produced at the earthed end of the tube, where beam spreading may become appreciable. An interesting consequence of the electrons' travel up the tube is the production of more secondaries by multiplication at the various electrodes, which would modify the potential distribution. It is desirable that the latter should be determined only by the external corona gaps fitted to the electrodes, and not by the internal currents.

To measure the total x-ray output, a calibrated gold-leaf electroscope was used (5·6 divs./min. $\equiv 10^{-5}$ r./sec.) of dimensions 3·3 × 3·3 × 5 cm. It was shown first that all the increased leakage was due to radiation formed in the presence of a beam: switching off the latter with maintained voltage always reduced the discharge to a very small value (natural leakage rate). The electroscope was set up 12 ft. from the tube opposite the high potential end, and 4 ft. below the level of the first tube gap. The results are shown summarized in table 1. The brass electroscope case was 1 mm. thick.

Table 1

Experimental condition	Rate of electroscope discharge (divs./min.)
Natural leakage, etc.	0·08
γ rays from 0·5 mgm. Ra 3″ from centre of electroscope	1·3
16 μa. beam ; 240 kv. ; tube current 0·6 ma.	1·2
20 ,, ; 240 kv. ; ,, ,, 0·6 ,,	2·0
4·5 ,, ; 420 kv. ; ,, ,, 0·3 ,,	5·2
3·7 ,, ; 300 kv. ; ,, ,, 0·3 ,,	1·7
Electroscope mounted 3″ from target	
3·7 μa. beam ; 300 kv. ; tube current 0·3 ma.	1·7

The safe dose is taken as 10^{-5} r./sec. (Kaye *et al.*, 1939) for 200 working hours a month. This corresponds to an electroscope discharge rate of 5·6 div./min. With the tube running at its maximum dissipation, say 160 μa. beam at 600 kv. with a total tube current of 1·2 ma., the safe discharge would be considerably exceeded unless the observers were shielded. No measurements were taken in these conditions.

Some performance figures for the tube without the separate focusing stage (see figure 1) are:—34 μa. target current for total tube current of 0·45 ma.

(positive and negative current) at 310 kv. with the ion source running at 9 kv., 0·5 ma., i.e. total canal-ray beam of $\sim$ 100 μa. The tube efficiency, relative to the collection of positive ions only, is good (probably $\sim$50 % or more). The inclusion of a collimator at the entrance of the tube gave no appreciable improvement.

The establishment of a weak magnetic field ($<$50 gauss) near the pump connection at the base of the tube gave a rise in accelerating voltage, for a constant van de Graaff generator output of 0·4 ma., from 390 to 450 kv., with a simultaneous rise in target current from 20 to 25 μa. Clearly the return current of secondary electrons had been much reduced. Target currents of 50 μa. were easily obtained. The tube voltage was reduced to 100 kv. to reduce the accelerating field for the secondaries, and we then obtained target currents of 8 μa. with no field and 3 μa. with the lowest tube gap shorted to reduce the field still further. In the latter condition 20 μa. were obtained with the field established near the pump connection. The secondary electrons were expected to arise mainly in the pump connection where the beam emerges from the last gap and impinges on a molybdenum diaphragm before entering the target tube.

It must be emphasized that the increase in target currents, as distinct from the rise in tube voltage for given input current when the secondary electron-deflector field is switched on, is largely due to a change in the tube-potential distribution. The voltage rise is due mainly to the decreased drain on the van de Graaff generator, which runs as a constant-current device.

Some results on potential distribution were obtained. A sphere gap comprising copper spheres of 6·25 cm. diameter on threaded shanks was placed across each tube gap in turn and adjusted with a long insulated rod to spark-over occasionally, but regularly, with the tube in operation. It was assumed that the presence of such a device would not affect the potential distribution: in view of the absence of corona discharge from the spheres this is probably justified. The sphere gap was mounted on the opposite side of the tube from the voltage-control corona gaps, which were not altered during a run. It was expected that tube surges would vitiate the results, but as they occurred at random they could usually be detected. The results are shown in figure 5. The total tube potential for (A) was 440 to 445 kv., and the sum of the separately measured gap voltages 448 kv. For (B) the respective figures were 445 and 569 kv. (surges might be expected to be serious here because of the appreciable beam current). For (C) the readings were 445 and 437 kv. The presence of a beam clearly altered the potential distribution, which was nearly linear in the absence of a beam. Figure 6 shows the voltages for each separate gap. The increased gradient at the low-potential end of the tube is the most prominent feature of the results. For all these tests, 1 in. long copper wires were attached to each electrode, pointing towards the positive electrode of each pair. This arrangement gives the most stable corona condition (confirmed by van Atta and Nowak, 1941). The corona wire on the earthed electrode brushed violently when large beam-currents were used, indicating (see figures 5 and 6) a large gap voltage. The total tube-current, ignoring small differences in corona current to earth from the various electrodes, will be constant at all points, i.e. a low internal electrode-to-electrode current necessitates a high external corona current for balance.

Several explanations of the effect of the beam on voltage distribution are possible. For instance, the cascading of secondary emission, i.e. successive multiplication at the different electrodes, would tend to decrease the gradient

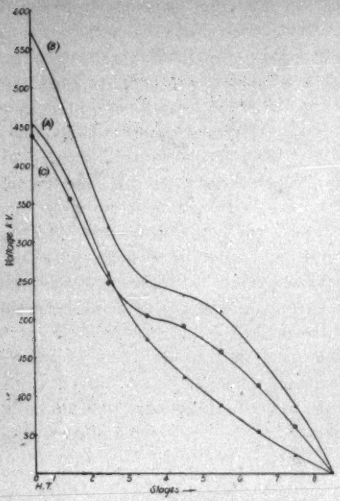


Figure 5. Potential distribution in cascade acceleration tube.

(A) No beam.
(B) 17·5 μa. beam.
(C) 5 μa. beam.

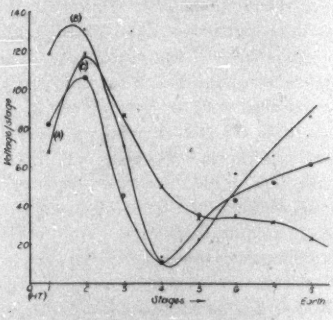


Figure 6. Variation of gap voltage along cascade acceleration tube.

(A) No beam.
(B) 17·5 μa. beam.
(C) 5 μa. beam.

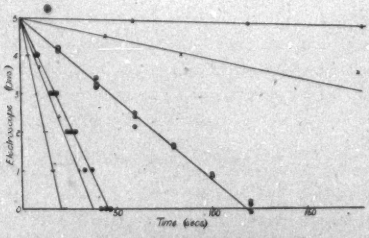


Figure 7. X-radiation from cascade tube.

* Natural leak.
· No absorber.
⊙ 0″·032 Pb.
× 0″·070 Pb.
⊕ 0″·0625 brass + 0″·25 glass.
\+ 0·2 mgm. Ra. 1″ from centre of electroscope.
0·98 ma. van de Graaff output.
520 kv. 2·5 μa.
Ion source :—5 kv., 1ma.

at the high-potential end of the tube. Again, the incoming ion-beam may lose its slow components to the upper electrodes: this suggestion is rendered unlikely by the known large effect of small deflecting fields at the pump connection, and the fact that inclusion of a collimator below the focusing gap had little effect.

The first tube electrode (highest potential) always became hot during operation, so it was clear that the secondaries had attained considerable energy. Probably many of the secondaries (~60 %) arose at the earthed end of the tube and travelled through several gaps before striking any of the electrodes. In view of the interest and importance attached to the problem, we decided to measure the hardness of the γ radiation from the tube, which should give information on the speed of the secondary electrons and thus on their region of formation. It was not expected that clearly defined results would be obtained, because of the continuous nature of the secondary-electron effects in the tube.

The electroscope previously mentioned was used for these tests. Some of the results are shown in figure 7. We attempted to correct for the absorption in the electrodes and glass tube walls by finding the effect of enclosing the electroscope in a similar electrode and glass tube, the equivalent lead thickness of which was taken as $0{\cdot}3 \pm 0{\cdot}1$ mm. ($\simeq 2{\cdot}75$ mm. copper).

In order to determine the hardness of the rays from the absorption data we have

$$\frac{I_1}{I_2} = \frac{e^{-\mu z}}{e^{-\mu y}} \quad \text{or} \quad \log_e \frac{I_1}{I_2} = \mu(y-x),$$

where I_1 and I_2 are the emergent intensities of radiation for equal incident intensity and for absorber thicknesses x and y respectively. It is assumed for convenience that μ, the linear absorption coefficient, is independent of x or y, i.e. that the radiation is monochromatic. This assumption is false in our case, but the results shown in table 2 indicate that the errors so introduced are not serious.

Table 2

Thickness of Pb absorber (mm.) excluding zero correction	μ	Equivalent speed of electrons (kv.)
3·43 to 1·78	10·4	250
1·78 to 0·81	7·8	280
1·78 to 0·81	13·5	230
1·78 to 0·3	16·5	210
1·78 to 0·3	18·5	190

The voltages corresponding to different values of μ are taken from Wintz and Rump (1931). The first reading is probably the most, and the last the least, accurate because of uncertainties attached to the zero correction. The deflection magnet mentioned above was not used in these experiments. If the radiation is homogeneous, i.e. if the differences in column 3 of table 2 are due entirely to experimental errors, the bulk of the secondaries must arise at the middle of the tube.

As mentioned above, it is remarkable that so little attention has been directed to the problem of secondary currents in acceleration tubes. For D.C. generators using thermionic valves, the increased tube current is usually not a disadvantage except for the heating and the x radiation emitted, but with electrostatic generators the problem is of prime importance (see § 2). Limiting diaphragms in the electrodes have been used since the early work of Cockroft, but it is doubtful

whether they were of great value. We made the necessary modified electrodes but were not able to test them.

Some further tests on focusing were carried out, to reduce the beam spread, which in turn would decrease the amount of secondary emission. Decreasing the corona gap on the lowest stages of the tube from 5 in. (point/plane) to 2½ in. (point/point) decreased the target current in a particular experiment from 70 to 22 μa. A remote-controlled corona gap was fitted to the highest potential stage but had only a small effect on focusing. On the second stage the effect was not observable. The separate focusing stage (figure 1) has a large effect on the tube performance, and by the use of it the target current may be almost doubled.

Finally, the use of a magnetic field of several hundred gauss in the tube near the pump connection resulted in a fall in the requisite input current to give 60 μa. target current at 450 kv. from 1·0 ma. with no field to 0·6 ma. with field. The effect of the secondary emission on the target current was negligible, i.e. about 10 %, shown by a separate experiment. The target was effectively shielded from the electrostatic field of the acceleration tube by the interposition of the tap and target tube, equivalent to a field-free region about 12 in. to 18 in. long and 2 in. diameter. A magnet was specially built for this work, and was installed below the earthed end of the tube. The best tube performance thus obtained was 160 μa. target current for 1·8 ma. van de Graaff current at 500 to 600 kv. The focusing voltage was 10 kv. for a source voltage of 6 kv. at 2 ma. current. The total positive-ion current entering the accelerating tube was about 200 μa. The tube-focusing efficiency is high and the secondary currents, amounting to about 1·5 ma., allowing for corona losses, are caused by ~50 μa. of lost ions, together with some neutral particles. Hence the magnification due to secondary emission at the electrode, or ionization processes in the gas of the tube, is about 30.

The beam current in the tube, with hydrogen, consists of about 50 % protons and 50 % molecular ions. A separate communication (Craggs, 1942) deals with this matter more fully, but the importance of the data in relation to the formation of secondary currents is easily seen. Measurements of secondary-emission ratios for H_1^+ and H_2^+ (78 to 426 kv.) show (Hill *et al.*, 1939) that values of about 6 and 3 for H_2^+ and H_1^+ ions respectively may be obtained for non-degassed copper surfaces. The molecular ions may be considered for bombardment purposes as two loosely bound atomic particles of half the equivalent H_2^+ voltage. Appreciable currents (Craggs, 1942) of fast neutral particles exist in the tube, so that a magnification of 30 or more may easily be obtained, owing to the almost certain existence of repeated secondary emission in the tube. In general, secondary-emission ratios tend to be approximately 2 for degassed copper surfaces (Allen, 1939), but the argument remains substantially unaffected.

§ 7. EXPERIMENTS IN PROTON BOMBARDMENT OF LIGHT ELEMENTS

It was decided that the first nuclear experiments should be on the γ-ray resonances of light nuclei. These have been investigated many times, but the absolute yields had not been accurately measured when our work was carried

out. Since that time some data have appeared (van Allen and Smith, 1941: Streib *et al.*, 1941). Another reason for the value of such experiments is that they provide a ready and accurate means of calibrating the apparatus and checking its general performance. The thin target resonances are often very sharp, e.g. the 440 kv. Li resonance is <11 kv. wide. Experimental data are given by Bothe and Gentner (1937), Hafstad and Tuve (1935) and Curran *et al.* (1938), besides the references already cited.

The γ rays in our case were directly detected with Geiger counters (Craggs and Smee, 1941) shielded from the intense x radiation from the tube by about 2·5 cm. of lead. The absorption of γ rays in the shield is appreciable (Crane and Lauritsen, 1935). The Li radiation intensity is reduced by a factor of approximately 4 for 1 in. Pb. However, in view of the overwhelming amount of x radiation, a thick absorber is necessary. Few experimenters use <1 cm. Pb.

The targets used are described in table 3.

Table 3

Bombarded element	Target
Li	LiCl fused on to copper or molybdenum LiOH fused on to molybdenum Li evaporated *in vacuo* on to copper Li_2O vapour condensed on copper
F	CaF_2 fused on to molybdenum
B, Be	Powdered and hammered into thin sheet lead
C	Aquadag baked on to copper

The stability of the counter was checked with a standard 0·2 mgm. radium source, 5 in. from the centre of the counter, before and after each run. Runs were taken with clean targets to assess the background count. The measured γ-ray intensities were never high, owing to the absorption in the counter shield and the small target current used. The C, B and Be yields were small compared with those from F and Li, and will not be discussed.

Representative runs are shown in figures 8 and 9. The focusing voltage was about 10 kv., to be added to the abscissae of the figures. Figure 9 was taken with F-contaminated LiCl. The voltage calibration of the apparatus was confirmed satisfactorily, but from time to time we noticed large deviations, due possibly to spurious charges on the corona screens affecting the generating voltmeter. We were not able to work with thin targets, because of the interruptions caused by more urgent work, but had developed the necessary evaporation technique.

Bothe and Gentner (1937) found relative yields from LiOH, CaF_2 and B of 30, 5, 1 at about 500 kv. Our figures for this voltage were 2·7, 16, 1. The possible errors were approximately ± 10 %, ± 10 % and ± 30 % respectively. We found the yields from CaF_2 to be consistently higher than those from LiOH, LiCl or Li_2O. The relative yields for Li, Li_2O, LiOH and LiCl were 8·5, 6·0, 3·6, 3·9. The discrepancies cannot be entirely due to experimental

errors (about ±10%). No correction has been made for absorption in the counter shield for either the Li or F radiations. The target solid-angles were maintained constant for all targets, and beam-size variations should be negligible. All yield figures given above refer to 500 kv. tube voltage.

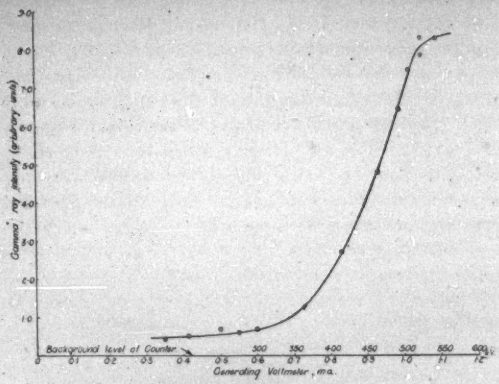
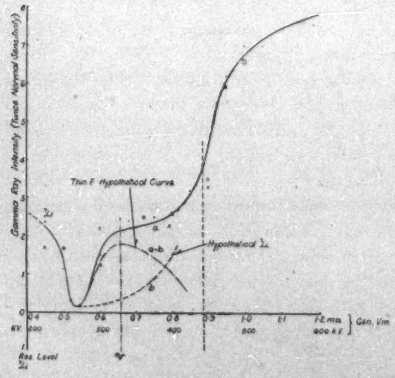


Figure 8. Evaporated lithium metal target.
30 μa. beam 0·2 Ra=10·5 γ-ray units.

Fowler *et al.* (1938) give comparable data. They stated the ^{11}B radiation (180 kv.) to be much weaker than that from ^{19}F (330 kv.) or ^{7}Li (440 kv.). Corrected yields for LiOH and CaF_2 were respectively 7×10^{-9} and 5×10^{-9}



Li Cl target (fused in molybdenum pan)
20 μa. beam
0·2 mgm. Ra ≃ 2×9·1 units

Figure 9. Gamma emission from fluorine-contaminated Li.

quanta/proton. The CaF_2 yield was therefore higher than that found by Bothe and Gentner, using Geiger counters and 1·5 cm. Pb. shielding. The latter (1939) give the relative counter response for the Li and B rays as approximately unity.

Hafstad and Tuve (1935) gave the Li-p output as equivalent approximately to 0·02 mgm. Ra/μa. at 500 kv. We performed an experiment with 0·2 mgm. Ra in a capsule (see above) fixed to the target, the 1-in. lead shield being removed from the counter, and obtained a count of about 4000/30 sec. LiCl bombarded with 35 μa. of mixed H_1^+ and H_2^+ ($\sim$50 % H_1^+) gave a count of 960/30 sec. at 550 kv., equivalent to 0·003 mgm. Ra/μa., i.e. approximately 0·007 mgm. Ra/μa. for an Li target. If we correct for the absorption of the Li γ rays in the counter shield, the agreement with Hafstad and Tuve data is reasonably good. More modern data for the F-p reaction (Streib *et al.*, 1941; van Allen and Smith, 1941) give the yield at 300 kv. as approximately $1{\cdot}4\times10^{-8}$ γ rays/proton, i.e. 9×10^4 γ rays/μa. sec., or about $1{\cdot}5\times10^{-3}$ mgm. Ra/μa. Unfortunately an accurate comparison between Li and F yields was not made by these authors.

In order to obtain more information on this matter we briefly investigated the absolute sensitivity of the Geiger counters, a knowledge of which might eliminate the direct calibration against radium. It is known that the γ radiation from light elements, in particular F (van Allen and Smith, 1941), is emitted equally in all directions from our target; hence a knowledge of the target solid-angle subtended at the counter, and the absolute sensitivity of the latter for the radiation considered, enable absolute yields to be determined. The method of Street and Woodward (1934) for calibrating cosmic-ray counters seems the most suitable for use with high-energy γ rays, but, due to lack of time, we were not able to perform similar experiments. The situation with regard to absolute efficiencies of γ-ray counters has been briefly outlined (Craggs, 1941), summarizing our own and other authors' results (e.g. Dunworth, 1940). For $\sim$2 Mev. γ rays the absolute efficiency of a normal counter is about 0·3 to 1 % and for our counters $\sim$0·5 %. The target solid-angle was estimated from counting experiments with radium (see above) to be about $0{\cdot}7\times10^{-2}\pi$. The count obtained with a CaF_2 target (5 μa. protons) at about 330 kv. was 520/20 sec. Assuming a counter efficiency of 1 per cent, the total count for a solid angle 4π was therefore 3×10^5/μa. sec., higher by a factor of 3 than the value of van Allen and Smith (1941). The error in assuming 1 % counter efficiency from radium measurements may be high enough to account for the discrepancy. As stated above, we were unable to carry out further experiments which, we hoped, would have given results of far greater accuracy.

§8. CONCLUSIONS

The development of a high-voltage apparatus for atomic disintegration work presents many additional problems when the available high-tension current is limited. These aspects of the technique have been specially emphasized, as it is believed they were not fully appreciated in other work.

Our experiments, unfortunately, are incomplete owing to the existing special circumstances, but target currents of 100 to 200 μa. have been obtained with 1·8 ma. high-tension generator current. This represents a good performance, despite the extremely low efficiency of the apparatus. When operated as a neutron generator, the set should give an activity of about 100 curies (Rn–Be), based on the data given by Heyn (1938): this is adequate for most purposes.

The original programme of work included certain experiments with neutrons—in particular, development work on neutron-sensitive Geiger counters and on various nuclear γ-ray effects.

§9. ACKNOWLEDGEMENTS

The author wishes to thank his colleague, J. F. Smee, for collaborating in some of the experiments described above. The work was carried out under the direction of Professor C. D. Ellis, F.R.S., to whom the author expresses gratitude for constant advice and encouragement. The London University Grants Committee provided means for purchasing the apparatus for ultimate use in King's College, London. However, many of the incidental expenses have been borne by the Metropolitan-Vickers Electrical Co., Ltd. Thanks should also be expressed to the Department of Scientific and Industrial Research for a grant which enabled part of the above work to be carried out; to Dr. T. E. Allibone, in whose laboratory the apparatus was developed, for his interest and suggestions; and to Dr. A. P. M. Fleming, Director of Research, Metropolitan-Vickers Electrical Co., Ltd., for permission to publish this paper.

REFERENCES

ALLEN, J. S., 1939. *Phys. Rev.* **55**, 336.
BOTHE, W. and GENTNER, W., 1939. *Z. Phys.* **92**, 45.
BOUWERS, A., HEYN, F. A. and KUNTKE, A., 1937. *Physics*, **4**, 153.
CRAGGS, J. D., 1941. *Nature, Lond.*, **148**, 661.
CRAGGS, J. D., 1942. *Proc. Phys. Soc.* **52**, 245.
CRAGGS, J. D. and SMEE, J. F., 1942. *Proc. Phys. Soc.* **52**, 245.
CRANE, H. R. and LAURITSEN, C. C., 1935. *Int. Conf. Phys. Report* (*Phys. Soc.*), **1**, 130.
CURRAN, S. C., DEE, P. I. and PETRŽÍLKA, V., 1938. *Proc. Roy. Soc.* A, **169**, 269.
DUNWORTH, J. V., 1940 *Rev. Sci. Instrum.* **11**, 167.
FOWLER, W. A., GAERTTNER, E. R. and LAURITSEN, C. C., 1938. *Phys. Rev.* **53**, 628.
GOODLET, B. L., EDWARDS, F. S. and PERRY, F. R., 1931. *J. Instn. Elect. Engrs*, **69**, 695.
HAFSTAD, L. R. and TUVE, M. A., 1935. *Phys. Rev.* **48**, 306.
HAILER, C., 1938. *Wiss. Veröff a.d. Siemens-Werken*, **17**, 115.
HERB, R. G., PARKINSON, D. B. and KERST, D. W., 1937. *Phys. Rev.* **51**, 75.
HILL, A. G., BUECHNER, W. W., CLARK, J. S. and FISK, J. B., 1939. *Phys. Rev.* **55**, 463.
KAYE, G. W. C., BELL, G. E., BINKS, W. and PERRY, W. E., 1940. *Rep. Progr. Phys.* **6**, 95.
KIRKPATRICK, P,, 1932. *Rev. Sci. Instrum.* **3**, 430.
KOSSEL, W., 1938. *Z. Phys.* **3**, 264.
LAURITSEN, T., LAURITSEN, C. C. and FOWLER, W. A., 1941. *Phys. Rev.* **59**, 241.
NEUBERT, V., 1938. *Z. Phys.* **110**, 334.
PEEK, F. W., 1929. *Dielectric Phenomena in High Voltage Engineering* (McGraw-Hill).
STREET, J. C. and WOODWARD, R. H. 1934. *Phys. Rev.* **46**, 1029.
STREIB, J. F., FOWLER, W. A. and LAURITSEN, C. C., 1941. *Phys. Rev.* **59**, 253.
TRUMP, J. G. and VAN DE GRAFF, R. J., 1937. *J. Appl. Phys.* **8**, 602.
VAN ALLEN, J. A. and SMITH, N. A., 1941. *Phys. Rev.* **59**, 501.
VAN ATTA, L. C. and NOWAK, W. B., 1941. *Rev. Sci. Instrum.* **12**, 305.
VAN ATTA, L. C., NORTHRUP, D. L., VAN ATTA, C. M. and VAN DE GRAAFF, R. J., 1936. *Phys. Rev.* **49**, 761.
WINTZ, H. and RUMP, W., 1931. *League of Nations* (*Health Organization*).
ZINN, W. H., 1937. *Phys. Rev.* **52**, 655.

A MACHINE FOR THE COMPUTATION OF STRUCTURE FACTORS

By R. C. EVANS, M.A., Ph.D., B.Sc.
AND
H. S. PEISER, B.A.,
Department of Mineralogy and Petrology, Cambridge

MS. received 30 *December* 1941 ; *in revised form* 15 *April* 1942

ABSTRACT. A machine is described which evaluates mechanically the expression $2f\cos(hx+ky)$, and its application to the computation of structure factors is discussed.

§ 1. INTRODUCTION

During the course of an X-ray analysis of a crystal structure, occasion repeatedly arises to calculate the structure factor of a proposed atomic arrangement, both when determining the signs to be attributed to the amplitude of each reflection, in order that Fourier syntheses may be prepared, and when confirming the proposed structure by an extensive comparison between observed and calculated intensities. Although the calculation of the structure factor is a very simple operation, it becomes tedious and time-consuming when repeated for a large number of reflections. The machine here described is designed to effect these calculations mechanically, and it has been found that under favourable conditions a very substantial saving in time is achieved by its use. In its present form, owing to lack of supplies of alternative materials, the machine is built almost exclusively of standard Meccano parts; it is, nevertheless, robustly constructed, and it has been found possible to realize a surprisingly high degree of accuracy, adequate for all the normal requirements of crystal-structure analysis.

The machine is primarily designed to derive the contribution F_{hk0} to the structure factor of the $(hk0)$ reflection due to two atoms, of atomic scattering factor f, situated at the points $\pm(x, y, -)$ and, therefore, related by an effective centre of symmetry at the origin $(0, 0, -)$. That is to say, the machine evaluates the expression

$$F_{hk0} = 2f\cos(hx+ky),$$

in which the co-ordinates x and y are expressed in degrees as fractions of the corresponding translations in the unit cell. The structure factor is required explicitly in this form in many analyses, but, as will be shown below, the machine can also be used to compute structure factors in the most general case. The structure factor of the structure as a whole, when required, must in all cases be obtained by summing the contributions computed by the machine for each pair of atoms.

§ 2. PRINCIPLES OF DESIGN

The mechanical principles of the machine are illustrated purely schematically in figure 1. The machine is operated by handles (A) and (B) which are turned

through a number of revolutions proportional to h and k respectively. The multiplication of h and k by x and y is achieved by variable friction drives between discs (C) and (D) and wheels (E) and (F), the displacements of which from the centres of the discs are proportional to x and y respectively. The angular displacements of the shafts (G) and (H) are thus proportional to hx and ky respectively, and these displacements are added in a differential gearing (I), the drive from the cage of which is so arranged that the rotation of a calibrated drum (J) and a crank (K) is $(hx+ky)$ degrees. A carriage (L) is held against this crank and is constrained to move in one dimension: its displacement is therefore proportional to $\cos(hx+ky)$, the resolved part along its direction of motion of the rotation of the crank. Finally, the multiplication of $\cos(hx+ky)$ by the variable parameter $2f$ is achieved by a pointer (M), pivoted about a fixed point on the carriage and sliding through a fulcrum (N), the deflection of the end of which is observed on the scale (O). The amplitude of this deflection can be altered by displacing (N) in a direction perpendicular to the length of the scale, and in this way variations in $2f$ can be accommodated.

§ 3. DESCRIPTION

The machine is illustrated in figures 2, 3 and 4.

The values of h and k are introduced by two handles (1), the drives from which are transmitted to steel-faced plywood discs (5) whose rotation is restricted to integral numbers of complete revolutions by the location on heart-shaped cams (6) of rollers mounted on spring-loaded jockeys (22). The values of h and k up to $\pm$ 20 are recorded on graduated drums (23) which can be reset to zero without the necessity of rotating the handles (1).

The discs (5) are spring-loaded against brass wheels (2) having sharp rims and carried on shafts (4) co-planar with and perpendicular to the axes of the discs. These wheels can be locked to their shafts (4) in any position, the displacement from the centres of the discs (5) being determined by the parameters x and y, the values of which are read on the scales (3). Accurate settings are achieved by the screws (20), whose hardened conical points form the outer bearings of the shafts (4); by rotating these screws the shafts, and with them the wheels (2), are displaced through small distances.

By relieving the spring pressures of the friction drives by means of the handles (8), the shafts (4), and with them the rest of the mechanism, can be operated independently of the handles (1) by turning the wheels (21).

The angular displacements of the shafts (4) are added together in the differential gearing (28), the rotation of which is transferred to a graduated drum (7) and to a crank (18) whose angular displacements are arranged to be $(hx+ky)$ degrees.

The carriage carries on its underside three parallel rails (27, 37 and 35) which rest on three wheels (32, 36 and 33) of a lower carriage. These wheels in their turn rest on fixed rails (31, 39 and 19), and the displacement of the carriage is geometrically constrained to motion in a straight line. A bar (34) on the carriage is pressed against a roller (30) on the crank (18) by the action of a spring (29), and the displacement of the carriage is therefore proportional to $\cos(hx+ky)$.

A pointer (10), pivoted at its centre of gravity on the shaft (17), is attached to the carriage and slides through a block (16) which acts as its fulcrum. At its

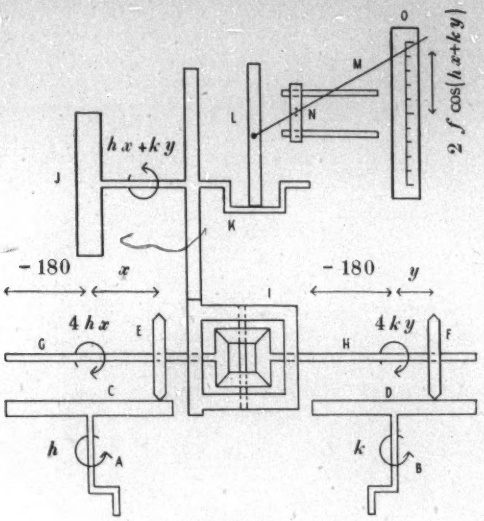


Figure 1.

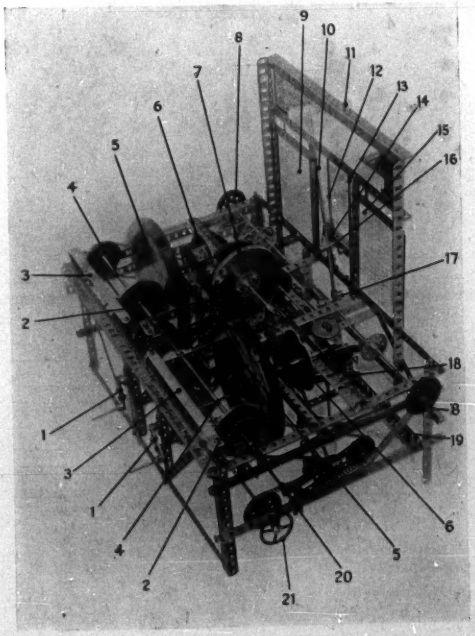


Figure 2.

To face page 458

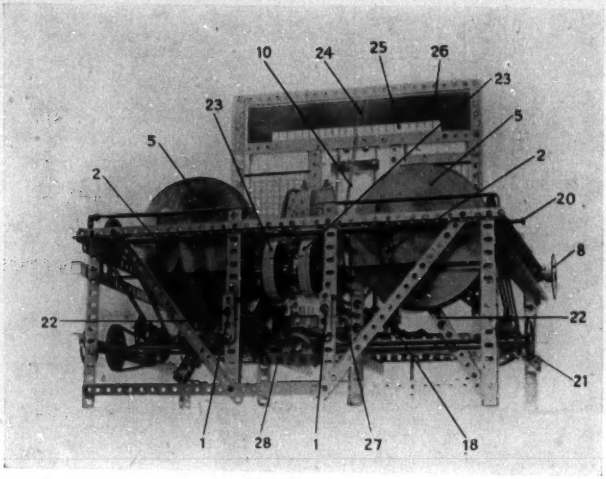


Figure 3.

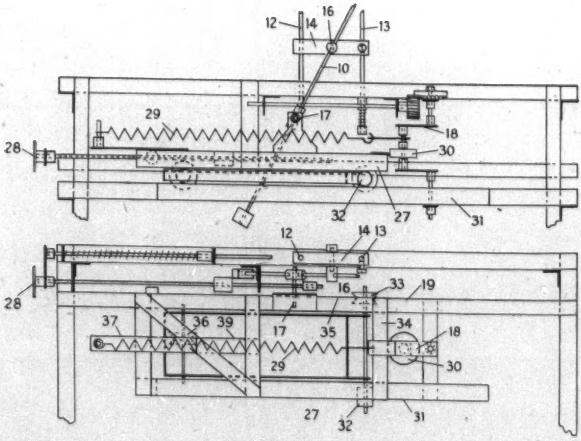


Figure 4.

upper end the pointer carries a thin needle (24) moving over a scale (25) fixed in front of a mirror (26) to enable readings free from parallax to be taken. The height of the fulcrum (16) above the pivot (17) can be adjusted by sliding the block (14) on two vertical guides (12 and 13). A fine adjustment of this setting is provided by a nut (11), and is facilitated by a handle (28) which displaces the carriage into the position corresponding to $\cos(hx+ky)=1$, so that the value of $2f$ may be read directly on the scale (25) without disturbing the handles (1).

Back-lash has been practically eliminated by spring-loaded gear wheels and by means of suitable jockeys for the chain drives. Care has been taken to reduce friction and inertia, especially beyond the friction drive, all danger of slip in which is eliminated by a 2 : 1 step-down between the handles (1) and the discs (5). The force required to drive the machine when applied to the rims of the wheels (2) is about 30 gm. weight.

§ 4. MANIPULATION

The values of x, y and $2f$ are introduced by the adjustment of the wheels (2) and the fulcrum (16). The scale (7) and the drums (23) are set to zero and the machine is ready for use.

The structure factor F_{hk0} for the $(hk0)$ reflection is displayed on the scale (25) after turning the handles (1) to the appropriate values of h and k; these values are recorded on the drums (23). Unnecessary effort in taking the readings may be avoided if the planes are arranged in a convenient sequence. The value of $(hx+ky)$ may be checked regularly, and if necessary corrected, by a direct comparison between the values of $(hx+ky)$ recorded on the scale (7) for the (500), (10, 00), (050), (0, 10. 0), etc., reflections and those given in the tables (9 and 15) which record the values of $5x$, $10x$, $15x$ and $20x$. In practice it is rarely found necessary to make any such correction unless the calculations are extended over several hundred reflections involving, in some cases, a total displacement of the scale (7) of some tens of thousands of degrees.

In many cases it will be unnecessary to make any allowance for changes of f with $(\sin\theta)/\lambda$. If, however, it is desired to make such an allowance, the computation may be interrupted at any stage and the value of $2f$ readjusted. In such cases it may be convenient to carry out the calculation in such a sequence that those reflections for which a single f value can be used are grouped together, since the time consumed in turning the handles (1) through a few revolutions is small compared with that required to reset $2f$.

In cases in which it is required to evaluate the structure factor F_{hkl} of the general plane (hkl), given by

$$F_{hkl} = 2f\cos(hx+ky+lz),$$

the calculations may be carried out as described above for the groups of planes $(hk0), \ldots (hkl), \ldots (hkL)$ successively. For this purpose it is only necessary to pre-set the scale (7) to the value of lz appropriate to each such group instead of to zero. If the number of reflections is large, the time spent in making these adjustments is quite insignificant. Values of F_{hkl} may be deduced from the relation $F_{hkl}=F_{\bar{h}\bar{k}\bar{l}}$ to avoid the necessity of setting the scale (7) to both lz and $-lz$.

Although the machine lends itself most readily to the computation of the structure-factor contribution due to two atoms related by a centre of symmetry, it can also be used in the general case of the space group 1. In this case the structure factor of a single atom is equal to f, but for several atoms the structure factor is given by

$$F_{hkl}^2 = \{\Sigma f \cos(hx+ky+lz)\}^2 + \{\Sigma f \sin(hx+ky+lz)\}^2$$
$$= \{\Sigma f \cos(hx+ky+lz)\}^2 + \{\Sigma f \cos(hx+ky+lz-90°)\}^2,$$

where the summation is carried out over all the atoms in the structure. In this expression each of the two terms in brackets may be evaluated by considering each atom in turn. For the latter term it is only necessary to set the scale (7) initially to 270° instead of to zero and then to proceed as before. The same preliminary setting of the parameters x, y and z will, of course, serve for both terms of the equation.

§ 5. EVALUATION OF STRUCTURE-FACTOR CONTRIBUTIONS OF ATOMS RELATED BY SYMMETRY OPERATIONS OTHER THAN A CENTRE

In those non-centrosymmetrical and centrosymmetrical space groups in which the numbers of equivalent general positions exceed one and two respectively, the machine can be employed to compute the structure-factor contribution due to each atom or pair of atoms separately. In general, however, much of the labour that would be involved in such calculations may be avoided. Thus, for example, in the space group $P2/m$, in which the structure factor is given by

$$F_{hkl} = 4f \cos(ky) \cos(hx+lz),$$

the calculations may be carried out successively for groups of planes with a common index k. For each such group, $4f\cos(ky)$ may be introduced as a constant quantity by the adjustment of the fulcrum (16), and if the number of planes for each value of k is large the saving in time achieved by the use of the machine is as great as in the simplest case of two atoms related by a centre. If however, the structure factors of planes of the type $(hk0)$ and $(0kl)$ alone are required, the advantage of using the machine is substantially less. Closely analogous arguments are applicable to all other space groups of the three classes of the monoclinic system, and will immediately suggest themselves on consideration of the relevant expression for the structure factor.

In many of the space groups of the orthorhombic system, the structure factor is in a form which lends itself less immediately to calculation by the machine. In all these space groups, however, there is, of course, at least one diad axis, and it is therefore always possible to apply to groups of atoms related by this axis arguments similar to those employed in the monoclinic system. In space groups of systems of higher symmetry, in which the number of equivalent positions is greater, the advantages of the machine are correspondingly reduced; in almost all cases, however, it may conveniently be employed to perform at least part of the calculations involved.

§6. RESULTS

Some typical results obtained by using the machine are recorded in table 1. For these observations the structure factors $9{\cdot}3\cos(h\,.\,154-k\,.\,78)$ and $9{\cdot}3\cos(h\,.\,32+k\,.\,172)$, corresponding to the arbitrarily chosen parameters (154°, −78°, −) and (32°, 172°, −), were computed for the sequence of 154 planes (000), (100) . . . (15, 00), (15, 10) . . . ($\bar{1}5$, 10), ($\bar{1}5$, 20) . . . (15, 20), (15, 30) . . . (15, 50) . . . (050), and the values obtained for the first and last and for two intermediate groups of five successive planes in this sequence are recorded. The calculated values of the structure factors, together with the observed and calculated values of $(hx+ky)$, are also shown in the table. From the figures it will be seen that the maximum error in F_{hk0} is 0·5—for (250)—and that this error is associated, as is to be expected, with a small value of F. The value of $(hx+ky)$ is in no case in error by more than 2°. An analysis of the complete set of 308 observations shows the errors in F to be distributed as follows:—

Table 1. Accuracy obtainable in a typical calculation

$hk0$	$h\,.\,154-k\,.\,78$		F_{hk0}		$h\,.\,32+k\,.\,172$		F_{hk0}	
	Obs.*	Calc.†	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
	°	°			°	°		
100	154	154	−8·3	−8·4	32	32	7·8	7·8
200	307	308	5·6	5·7	64	64	4·2	4·1
300	100	102	−1·7	−1·9	96	96	−0·8	−1·0
400	254	256	−2·6	−2·2	128	128	−5·6	−5·7
500	50	50	6·0	6·0	160	160	−8·9	−8·7
$\bar{5}20$	154	154	−8·3	−8·4	184	184	−9·3	−9·3
$\bar{4}20$	308	308	5·7	5·7	215	216	−7·7	−7·5
$\bar{3}20$	102	102	−1·9	−1·9	248	248	−3·7	−3·5
$\bar{2}20$	254	256	−2·6	−2·3	279	280	1·5	1·6
$\bar{1}20$	50	50	5·9	6·0	312	312	6·3	6·2
$\bar{5}40$	356	358	9·3	9·3	168	168	−9·1	−9·1
$\bar{4}40$	150	152	−8·0	−8·2	200	200	−8·8	−8·8
$\bar{3}30$	304	306	5·2	5·5	231	232	−5·9	−5·7
$\bar{2}40$	99	100	−1·3	−1·6	263	264	−1·0	−1·0
$\bar{1}40$	252	254	−3·0	−2·6	296	296	4·1	4·1
450	224	226	−6·8	−6·5	268	268	−0·4	−0·3
350	70	72	3·2	2·9	236	236	−5·2	−5·2
250	276	278	0·8	1·3	204	204	−8·6	−8·5
150	122	124	−4·9	−5·2	172	172	−9·2	−9·2
050	329	330	8·0	8·1	140	140	−7·1	−7·1

* i.e. obtained by means of the machine.
† i.e. calculated from $2f\cos(hx+ky)$.

Error	0·5	0·4	0·3	0·2	0·1	0·0
Frequency of occurrence	2	14	24	63	105	100

The accuracy of the machine is thus amply adequate for all normal requirements of structure analysis.

The time saved by the use of the machine was estimated by computing the structure factor of two groups of planes, containing about 300 and 600 reflections respectively, both with the aid of the machine and by numerical methods. For the latter calculations, standard tables of hx and ky, reduced to angles in the range 0 to 360°, were employed to give $(hx+ky)$. The value of $\cos(hx+ky)$ was taken directly from trigonometrical tables and the multiplication by f was carried out on two slide rules. The results of this comparison are shown in table 2. It is seen from these figures that the time saved by the machine is about 60 % of that required to make the calculations from tables. Where,

Table 2. Speed attained in a typical calculation

	Machine			Numerical methods		
	300 planes	600 planes		300 planes	600 planes
Setting x and y	10 min.	8 min.	Computing $(hk+ky)$	31 min.	70 min.
Recording F_{hk0}	31 ,,	86 ,,	Evaluating F_{hk0}	77 ,,	170 ,,
Total	41 ,,	94 ,,	Total	108 ,,	240 ,,

however, it is necessary to take account of variation of f with $(\sin\theta)/\lambda$ the advantage of using the machine is substantially reduced. On the other hand, much of the time spent in performing calculations with the machine is employed in recording the observations, and if an unskilled assistant is available to transcribe the results as they are dictated, a further substantial saving in time is achieved. Thus it was found that 600 observations could be recorded in this way in 41 minutes, compared with 86 minutes when the operator worked alone. A further 8 minutes was required for setting the x and y values, so that the total time for the observations was 49 minutes, compared with 240 minutes when tables were used.

In conclusion it may be pointed out that computations can be performed with less fatigue with the help of the machine than by numerical methods. The chance of accidental error is, moreover, reduced by its use.

§ 7. ACKNOWLEDGEMENTS

It is a great pleasure to express our thanks to Meccano, Ltd., and especially to Mr. F. Riley, B.Sc.,who made available to us the parts used in the machine, and to Mr. A. N. Lanham, who constructed those parts which could not conveniently be made of Meccano. One of us (H. S. P.) is indebted to the Council of St. John's College, Cambridge, for a Hutchinson Studentship.

REVIEWS OF BOOKS

The Nature of Thermodynamics, by P. W. BRIDGMAN. Pp. xii + 229. (Cambridge Mass.: Harvard University Press; London: Sir Humphrey Milford at the Oxford University Press, 1942.) $3.50. 20*s*. net.

In this striking book, Professor Bridgman analyses the concepts of thermodynamics from the " operational " point of view which he has described in previous writings and which maintains that the meaning of a physical concept is wholly expressed by the set of operations by which the concept is determined. Though there are still wide divergences of opinion on the matter, it is probable that the great majority of physicists have been much influenced by the operational outlook, and they will accordingly turn with interest to this minute analysis of the laws of thermodynamics which, as the author truly says, smell more of their human origin than other physical laws, and have also the distinction of being the only ones which have retained both their significance and their importance without appreciable change since the mid-nineteenth century.

It may be said at once that probably no such searching analysis of thermodynamic concepts has previously been made, and that anyone who can peruse this book carefully without clarification of his own ideas is either incredibly sagacious or essentially ineducable. It is the kind of book that should be read first fairly rapidly and then again, at least once, with more care. If the following remarks are almost entirely critical, that is because such treatment is the most fitting tribute that can be paid to a most stimulating attempt to bring order into the most elusive group of physical ideas.

The bulk of the discussion is contained in two chapters dealing with the first and second laws respectively ; two shorter chapters are given to miscellaneous questions. Throughout the book, Bridgman rightly insists that macroscopic are more fundamental than microscopic considerations, since all instrumental operations are necessarily macroscopic ; hence it is invalid to reject any conclusion properly drawn from thermodynamical laws simply because it does not fit into the present form of statistical or kinetic theory. This needs emphasizing, but it is unfortunate that—doubtless through an oversight—after the remark that " the task of thermodynamics is to give us mastery of what happens .... when temperature differences appear in the system ", the essence of thermodynamics is introduced in such an anticlimax as " A thesis of thermodynamics is that certain very broad aspects of the new and extended phenomena can be exhaustively handled by the introduction of the temperature of the various parts of the system ". It should be added that the succeeding analysis has infinitely more substance than this would lead us to expect.

The first law is expressed in the form, $\Delta E = \Delta W + \Delta Q$, and attention is concentrated on the operational meaning of the three terms in this equation. Bridgman recognizes that the actual use of the law goes beyond what can be strictly justified operationally, but he is less concerned to condemn this than to understand exactly what the actual use implies. He finds in it much " verbalising "—i. e. not merely " pencil and paper " operations, but also elements of the mental pictures which we associate with the concepts, not all of which have operational sanction. He views this, and also an occasional " mildly vicious circularity " of argument, with a leniency which will command much sympathy from physicists more interested in acquiring new territory than in establishing their right to it ; but, as events have shown more than once in the history of physics, logic is a jealous goddess, and indifference to her demands is liable to meet with retribution.

For example, in an interesting discussion of the localization of the energy " transmitted " from a motor through a belt to a flywheel, Bridgman introduces a Poynting vector to represent the mechanical energy flux, and later finds himself compelled to " say " (he admits that he is verbalizing) that the energy flux must be along the belt. " If instruments outside the belt give no readings ", he says, " then I have to *say* that the path of energy flow is in the belt ". It is a perilous argument. He later gives reasons for believing that there should be a field of force (obviously extremely weak) related to a

gravitational field in the same way as a magnetic is to an electric field, and so a gravitational Poynting energy vector. If such a field should accompany the stress in the belt, the fact that external instruments give no readings would be due merely to their insensitiveness, and the flow of energy might be directed, as with a wire carrying an electric current, from the surrounding medium into the sides of the belt. The point is that the *direction* of energy flow is independent of the *magnitude* of the energy components, and a verbalization which relates them invites disaster.

The whole treatment of the localization of energy, which forms the backbone of the discussion of the first law, seems to me to be based on an error. Modifying a view taken in an earlier book (*The Logic of Modern Physics*), Bridgman feels compelled to ascribe " full and complete " localization to energy because, he holds, the first law is applicable to any region, whether occupied by matter or not, defined by a surface which separates it from its surroundings. By drawing the surface in all possible places we are therefore forced to grant complete localization to energy. Here it seems that he has been betrayed into forgetting his own operationalism. He is at infinite pains to elucidate the operational procedure for measuring the ΔW and ΔQ that cross the boundary (even so, it is not easy to see how to apply the method of measuring ΔQ in empty space), but he nowhere touches on the impossibility of operationally identifying the boundary itself. That this is impossible in space unoccupied by matter is the plain meaning of the special theory of relativity, and if you cannot localize a surface, then clearly you cannot localize the energy inside it. Hence the whole argument for energy localization collapses.

Indeed, one is inclined to think that the greatest obstacle to Bridgman's quest for understanding is his barely disguised dislike of the viewpoint of relativity, notwithstanding that it was the confessed origin of his own operational method. His references to relativity constitute a sort of gesture of disapproval rather than an open attack, and he often loses the significance of the relativity outlook by regarding its requirements as merely a " small correction " which may be ignored. This is a fatal attitude to the problem of understanding physical concepts, reminiscent of that of the erring damsel who excused her misdemeanour on the ground that the baby was only a very little one. As a prominent example we may cite the remark, on which much is based, that " the mass of a stone is unchanged (except for relativity effects) when it is measured in a frame of reference moving with respect to the original frame, but its kinetic energy is drastically altered ". The fact is that the changes of mass and kinetic energy are precisely equivalent, but are represented by numbers of different magnitude simply because different units of measurement are chosen. One gram is c^2 ergs, and to say that the change of mass is negligible compared with the change of kinetic energy is tantamount to saying that the change in E arising from ΔQ is negligible compared with that arising from ΔW because there are 40 million ergs in a calorie. Incidentally, Bridgman nowhere mentions that the first law, in the form given, is true only when the units of measurement of ΔW and ΔQ are the same. As measured by the operational processes described, they would normally be obtained in very different units, and it is at least arguable that the essence of the first law lies in the fact that ΔW can be added to ΔQ.

The discussion of the second law is largely parallel to that of the first, with entropy occupying the place previously taken by energy. Thus the sense in which we can speak of a localization and a flow of entropy is examined, and limits are sought to the legitimacy of ascribing " physical reality " to entropy. There is a very suggestive discussion of the meaning of reversibility and recoverability, and the statistical aspect of entropy is considered at some length. A possible extension of the second law to cover irreversible phenomena is adumbrated ; according to this, the irreversibility of a process would be represented by a definite assignable increase of entropy, instead of an indefinite one as in the ordinary formulation of the law.

There is much in this chapter on which one would like to comment. One point must suffice, namely, the conclusion that black-body radiation accompanied by fall of temperature is reversible. This is certainly a startling result, which if true, is of fundamental importance. We give the argument in Bridgman's own words :—

> " Idealise the universe as a sphere at uniform temperature which has been prevented from radiating since its creation by an impervious shield. Remove this shield, allowing the sphere to radiate a finite ' gulp ' of radiation, and replace the

shield. The entropy of the sphere has decreased, but to compensate there is entropy traveling in the gulp of radiation. Next catch this gulp of radiation on a perfectly reflecting spherical mirror placed at the confines of the 'universe', reversing the direction of radiation. The radiation eventually returns to the sphere, is there absorbed, and the sphere in the end resumes its initial temperature. The initial entropy of the sphere is therefore also resumed. There is no other place where there may be any entropy, so that the net result of the emission and re-absorption of the gulp must be that the total entropy of the universe is unchanged. But since the entropy of a closed system can never decrease, this means that the total entropy was always the same. The entropy associated with the gulp of radiation while it was traveling through space was exactly equal to the entropy lost by the sphere. Furthermore, since the occurrence of an irreversible process leaves its indelible impress in a permanent increase of entropy, it must be that no irreversible or entropy-increasing processes occurred. That is, neither the emission of radiation by the sphere in the first place nor its re-absorption later is to be described as an irreversible process. This, I believe, is quite different from what would be the snap judgment of many physicists. It is evidently the unidirectional character of the radiation that distinguishes this special situation. Although the unidirectional emission of radiation is a process which occurs spontaneously and at a finite rate it is not an irreversible or entropy-increasing process, as is the conduction of heat down a temperature gradient, a process which to our first crude physical intuition might 'feel' to be the same."

Bridgman goes on to say :—

> "The entropy increase arising from .... exchange of radiation between bodies at different temperatures is not to be sought in the initial act of absorption, which may be non-entropy-increasing, but is to be found after the initial absorption in the spreading out of the spectrum of the absorbed energy from the distribution characteristic of the higher temperature of its source to the distribution characteristic of the lower temperature of the sink."

This argument does not appear to me to be sound. In the first place, the radiation is not "unidirectional"; it proceeds in all directions from the source. "Symmetrical" or "isotropic" is the word required. This is not merely a verbal difference, because whereas unidirectional radiation is unusual, and would justify the phrase "special situation", radiation outwards from the source is normal, and in this respect the artificial situation envisaged is in no way "special". But the chief criticism is that the process of removing the screen to re-admit the returning radiation would necessarily allow another "gulp" to escape, and the original situation would not be restored. And, in fact, in the nature of things it cannot be restored, for there is no conceivable way of restoring it without the use of a "shield" which transmits all radiation in one direction, but none in the opposite one. We have no experience of which this can be regarded as a legitimate extrapolation. We may admit a perfect reflector since we can realize it with close approximation, but a one-way transmitter is a sheer fantasy. Consequently, when the first gulp of radiation left the universe something physically irreversible happened, and the entropy must accordingly have increased.

There is another aspect of the passage quoted which is of considerable significance, namely, the ascription of the entropy increase when radiation is absorbed to the re-adjustment of the spectrum of the absorbed radiation, and not to the act of absorption itself. This shows a considerable departure from the "operationalism" of *The Logic of Modern Physics*, for how can one distinguish operationally between the initial absorption of radiation and the re-arrangement of its (obviously unobservable) spectrum while it remains absorbed? The distinction is purely verbal, and it is with surprise and some regret that we find Bridgman not only adding "paper and pencil" to "instrumental" operations, but also including the formation of a mental picture among the operations which give significance to a physical concept. Thus on pages 163–4 he writes :—

> "One might be tempted to say that 'operationally', since there is nothing to distinguish between the two views, there is in fact no difference, and that they must be regarded as the same. This would indeed be true, only 'operationally'

must cover *all* the operations, not merely those of numerical calculation and program drawing. In particular, the operations include the entire verbal structure into which the two views may be fitted. From this point of view there is a difference which must not be overlooked between the common-sense point of view which says that a pail of water has never frozen on the fire, and I shall make my programs for the future without leaving any room for this actually happening although I recognize it as a *possibility* that it *might* happen, and the point of view that says that there is a chance of 10^{-1000} that a pail of water might freeze on the fire but this chance is so small that it would be waste of time to figure what course of action I had better adopt if it did happen ".

One cannot help feeling that in thus enlarging the meaning of " operation ", Bridgman is losing the chief value of the operational point of view. As originally formulated, that point of view did make clear a sharp distinction between possible directions of physical speculation, but if operations are to include the " mythology " with which we find it comforting to gild the philosophic pill, it is difficult to see that any distinction at all is left. Even absolute simultaneity passes the operational test by virtue of its unique " verbal structure ". One may grant that an insistence on instrumental operations alone would prohibit progress along directions that might prove both legitimate and profitable. " Paper and pencil " operations, in the sense of strictly specified mathematical procedures, would seem to be necessary for the full development of physics. But Bridgman's pre-occupation with " physical reality ", which he admits surprises himself, indicates a shift of interest from the conceptions of physics to the misconceptions of physicists. That is a misfortune which it may be hoped is only temporary.

The possibility must be faced, however, that this *reductio ad absurdum*, as we might almost call it, is the inevitable conclusion of the rigorous development of the operational method, just as the logical positivists' original criterion of meaning has issued in realization of the nonsensical character of their own statements. It has always seemed to me that " operationalism " is a good servant but a bad master. It is excellent, for example, to realize that a measurement of a velocity in terms of the space covered in a given time is something different from a measurement in terms of the displacement of a spectrum line, and there are occasions (possibly the phenomenon of the " expanding universe " is an example) when recognition of the difference may be of cardinal importance. At the same time, the construction of a theory according to which the measurements are regarded as equivalent is a great step forward and is responsible for a great part of our present knowledge of astrophysics. A fundamental definition of physics demands, first of all, recognition of the distinction between the experiences which are our data and the rational concepts which we use to correlate them, and a doctrine such as operationalism which can be expressed only with difficulty in terms of this distinction is thereby marked as of doubtful fundamental validity. That is not to say that it is not of extreme value on its own level, and we end as we began by an acknowledgment that Professor Bridgman's study is one of the most valuable contributions to the understanding of thermodynamics that have yet appeared.

HERBERT DINGLE.

Science and Education (*Current Problems*, 15), by S. R. HUMBY and E. J. F. JAMES. Pp. viii + 146. (Cambridge University Press, 1942.) 3*s*. 6*d*.

It has become increasingly evident that there is a great gap between what science has done and what it might have done for society. This book asserts that the gap is due to the unawareness by society of the social changes produced by science in the past and, therefore, of the possibility of the solution of our present social problems by science. It traces this unawareness partly to the absence of science from the curriculum of the classical sides of many public and grammar schools and partly to the unsatisfactory nature of the little science which is taught to the other boys in these schools. It therefore makes radical constructive suggestions about the content of, and method of presentation in, the science courses in schools and universities.

The book commences with an outline of the social changes of the last three centuries due to science, citing examples such as the growth of urban populations, the new structure of society produced by the industrial revolution, and the possibility of the age of plenty.

It suggests that the evils associated with great cities (their smoke, their slums, their disease) and the continued existence of poverty in the face of the possibility of plenty can be eradicated by the application of scientific knowledge and scientific method. The failure to do so is attributed both to the fact that most of our cabinet ministers and civil servants learned little or no science at school, and also to an unawareness on the part of society generally of what science has done and therefore could do to influence social change. While the evidence for the gap between what science has done and what it might have done for society is good, the arguments that this gap is due to unawareness are less convincing. For example, the authors state that " this fatal division between resources and unawareness is shown in any number of ways in almost every department of our life ". The reader expects a selection of striking examples, only to find an apocryphal story about Gladstone and Faraday! It may be that there is good evidence ; if so, would it not have been wiser to quote it? And is the unawareness quite so glaring as the book suggests in view of the fact, quoted in the book, that science is now a profession, whereas a hundred years ago it was the pursuit of a few wealthy amateurs? And does not the creation during the last fifty years of the National Physical Laboratory, the Department of Scientific and Industrial Research, the research departments of the various trade associations and of many private firms show that society is not altogether unaware of the value of science to industry, at least?

Even if the explanation of the gap between the achievements and possibilities of science is not accepted, the suggestions made in the book about the modification of the content and presentation of science in schools and universities would help to reduce that gap and so deserve the most careful examination. The authors rightly condemn the paucity of time still devoted to science in public and grammar schools, and the prejudice which still exists against it ; this results both in a large proportion of the able boys being forced to do classics, and in a divorce between humane and scientific studies, the one being regarded as cultural and the other as merely utilitarian. The authors plead for a broader, more humane course of science for the future citizen (as distinct from the future scientist), which should include not only chemistry and physics, but also some biology, astronomy and geology, and should be linked with history by the study of the social changes produced by science. They emphasize the need to infuse the essentials of the scientific method into the future citizen, so that he may be able to distinguish truth from propaganda and be better able to weigh objectively and dispassionately the issues which a citizen must continually face. Finally, they recommend a course of general science of two periods a week after school certificate for *all* specialists, in which recent scientific knowledge and the application of science to modern social problems is discussed. They claim that if a future civil servant or cabinet minister had learned science in this way, he would know what type of social and political problem was capable of solution on scientific lines, and how much progress could be expected in a limited time ; he would, in fact, be able to judge when to call the scientist to his assistance.

The book is an interesting and critical discussion of a problem of vital and universal interest, and it should be read by university teachers and schoolmasters and should be in all university and school science libraries. A. W. B.

A History of Science, by Sir WILLIAM CECIL DAMPIER. Third Edition. Pp. xxiii + 574. (Cambridge University Press, 1942.) 25*s*. net.

The preface assures us that this is, in effect, a new book. It is at the same time a third edition and has the advantage which comes from revision and enlargement. But phrases like " in this year 1929 " do occur here and there, and suggest that it has not lost all traceable likeness to the second edition (1930).

The decade 1930 to 1940 is covered by the addition of an eleventh chapter, and as the work begins with the scientific thought of the ancient world it can be easily imagined what a stupendous task has been undertaken. But this in itself gives little idea of what has been attempted, for to the history of science is joined the development of philosophic ideas, which bulk quite largely ; and, as if this were not enough, considerable space is devoted to the discussion of religion. This last subject seems to conclude (p. 489) in the theory that ritual transcends dogma in importance. It is, perhaps, not altogether unfair to infer that the work is written from the angle of the college chapel, preferably with a good choir.

As it is quite impossible to cope at all adequately with an extremely interesting and suggestive work of such vast extent in a short review, it seems better to quote the author's own candid admission (p. 493), which can hardly be final: "Later views have brought into prominence older discoveries which at that date (1930) appeared of less interest. Hence, at some points, it is necessary to turn back in time and describe older investigations, ignored in the first two editions of this book". But a little thought may perhaps be given to a question which arises naturally in a work of this kind, the division of time into significant historical epochs.

To approach this question in a concrete form, how should the eighteenth century be considered? The choice is made in the belief that this period has been most commonly neglected. Now, if there is a period of scientific history which is clearly defined, it is surely the seventeenth century, the age of Galileo, Huygens and Newton. Moreover, as the age of Des Cartes, Spinoza and Leibniz, it is equally defined from the philosophic angle. Thus, for want of a better, the year 1727 may be suggested as the beginning of the next epoch. How are the following years to be regarded? From the English point of view there is a strong temptation to pass them over lightly. The English universities, with their collegiate structure, suffered from the ills to which their constitution is peculiarly prone, and cut a very sorry figure. But the eighteenth century in science cannot be dismissed lightly. It is the age of Euler and Lagrange, it carries electricity from the terrestrial magnetism of Halley to the days of Faraday, and science, with the addition of technology, is widened in its entire scope. All these things bring another question. If, as it appears, the eighteenth century represents something which deserves to be considered on its merits, what limit in time should be imposed? A Cambridge man often likes to think of a *risorgimento* occurring about the year 1820. But this parochial view may be corrected by trying to see the matter through French eyes. In spite of revolution in the political sphere, no discontinuity is visible in science. In mathematics the progress from Clairaut to Lagrange is maintained from Cauchy to Poincaré without any palpable break, and elsewhere a similar continuity can be traced. Surely a clue is beginning to emerge in a single age of the classical mechanics and the unbroken atom, starting from the death of Newton, embracing the use of technology and ending about 1895. If it must be separated into two halves, the Victorian age may serve for the second, but there is no organic interruption. Should it be objected that this means giving too much weight to the influence of the physical sciences, the answer is that the others must either keep step or the attempt to treat the whole in one synthesis is doomed to failure. As time passes, it must become more difficult to preserve an unbroken front.

This scheme, it must be confessed, fares ill at the hands of Sir William Dampier, and the divergence begins early. It flows from the treatment of the seventeenth century, which is divided at 1660. The result is to isolate the Newtonian epoch (chap. iv) from all that went before and clear the decks for the nineteenth century, which is discussed in three separate chapters (v–vii) followed by four more (viii–xi) on still more recent times. On the other side, after the second chapter has traced the Middle Ages from the decline of Rome to the end of the thirteenth century, the interval from scholasticism at its height, with science feeling its way step by step, is continued past Galileo to the foundation of the Royal Society. Such arrangements are doubtless a matter of free choice, and the individual ought not to be hampered. But, according to the test which has been suggested, the treatment accorded to the eighteenth century appears rather less than its due.

Scientific opinion in general has little sympathy with philosophical ideas. Like the curvature of an arc, philosophy can be ignored in science so long as the range is sufficiently small. In the distant past it intrudes and must be taken into account. It occupies considerable space in this book, and this is natural enough. That philosophy is once more claiming attention in the science of the day is, perhaps, merely a symptom of adjustment of ideas to that curvature which may have been too long ignored and resisted. But the philosophy of the future is likely to differ widely from that of the past because, to satisfy the needs of science, it must be evolved from within, not imposed from without. Like mathematics, where the point can no longer be contested, science (if worthy of the name) must be master in its own house.

From this point of view the eighteenth century refuses to be treated as a part of the Newtonian epoch. For it is in this age that the claim of science to independence from philosophy and religion is not only made but conceded, however grudgingly, for the first time. The limited supremacy of Newton can also be illustrated by considering

two individual astronomers, both of the first magnitude. James Bradley owes this eminence not to the discovery of aberration (the only achievement mentioned in this book), but to the fact that he belongs to that small class who have raised the general standard of accuracy in observation. He is much more closely allied to Flamsteed than to Newton. The other astronomer is Sir William Herschel, only mentioned here on account of an isolated experiment on the infra-red solar spectrum. In spite of his discovery of Uranus, his greatness lies in the fact that he looked far beyond the solar system, which was Newton's main concern, and enlarged the boundaries of the visible universe. Both these astronomers, in different ways, illustrate another characteristic of the age in which they lived, the growing importance of finer and larger instruments. But more than enough has been said on behalf of the eighteenth century to show that, even apart from the biological sciences, it cannot be fairly treated as a Newtonian preserve. H. C. P.

Introduction to Physical Chemistry, by ALEXANDER FINDLAY. 2nd Edition. Pp. vi+582. (London: Longmans, Green and Co., 1942.) 15*s*.

This second edition of a widely used book is some ninety pages longer than its predecessor. The new material is distributed fairly generally throughout: notice should be made of two new short sections on *Elements of Crystallography* and *Molecular Dimensions and Configurations*. The number of problems has also been increased. These additions result in some shifting of emphasis, and much of the discussion has been re-written and references to the literature have been brought up to date.

The presentation of any work of this kind is bound to reflect, in some way, the scientific pre-occupations of the author. It is not surprising, therefore, to find the chapter on the Phase Rule one of the best in the book.

As befits an introductory text-book, the discussion is general and descriptive, but it is, perhaps, a pity that there is so little of more precise treatment. Both kinetic theory and thermodynamics suffer somewhat in consequence. It is also unfortunate that (p. 293) a distinction is made between thermodynamics and thermochemistry, whereby the former has the now standard " acquisitive " sign-convention, the opposite convention being used for " thermochemistry ". Surely the former convention only is required.

The first edition provided a well-written introduction to the subject: the matter added to the new edition has brought it into line with recent developments, and increased its usefulness. R. F. B.

Biography of the Earth: Its Past, Present and Future, by George Gamow. Pp. xiv+242. (New York: The Viking Press; London agents: Macmillan and Co., Ltd., 1941.) 12*s*. 6*d*.

Professor Gamow has followed up his very successful *Birth and Death of the Sun* with a volume devoted to our own planet.

Starting from the evidence of radioactivity, the date of solidification of the oldest rocks is shown to be not more than two thousand million years ago (or two billion in the American terminology, which the author uses). The fact that it separated from the sun is next demonstrated, and the generally accepted hypothesis as to the nature of the close approach of two stars which led to this event is described.

Next comes the problem of how the moon was separated from the earth, and after that the question of how it got so far away. Then the arguments are marshalled for assuming that the Pacific Ocean represents the scar left when the moon was removed; the fact that the shape looks right is obvious to all of us, but the strongest single item of evidence lies in the fact that whereas everywhere else on the earth the basalt is covered by a granite layer 50 to 100 km. thick, yet in the Pacific Ocean no granite is found—the moon naturally took the whole of the top layer away.

More detailed consideration of the planets, their atmospheres, and the possibility of life on them brings us back to the detailed study of the earth's structure and interior. Here the hypothesis of continental drift is discussed, and isostasy is explained in a way which should be intelligible to the layman—no mean feat. This part of the book deals also of course with the processes of mountain building, ice ages and vulcanism. In

connection with ice ages, the author rejects rather summarily the theory, which is widely held by many authorities, that variations in the amount of carbon dioxide or water in the atmosphere, by changing the absorption coefficient for radiation, may have been responsible for major changes of climate. He attributes the whole of the effects to changes in the earth's rotation round the sun, without even mentioning any of the arguments against this theory.

Chapter IX deals with life on the earth, and in particular with its origin. Plausible arguments are advanced for the theory that life might arise spontaneously, and it would be interesting to have the opinions of a biologist on the matter.

The book closes, unlike most biographies, with a glimpse into the future. As far as can be foreseen at present, there will " soon " be a period of much more intense mountain building. Then, about 50,000 A.D., there will be an ice age, when most of Europe and North America will be under thick ice, which will cover Oslo, Copenhagen, Stockholm and Leningrad, but not London, Paris or Berlin. Before this, however, the climate of the earth as a whole will be warmer, with its climax about 20,000 A.D. Many millions of years later again, the mountains will have been washed flat. It is suggested that the mammals, all increased in size, will still dominate the animal kingdom.

While this is happening, the moon may be expected to recede from the earth, until in 20 or 30 thousand million years she reaches her maximum distance, after which she will gradually approach the earth again. At the culminating point, the day and the month will be equal, each occupying 47 of our present days. All this, however, depends on the sun remaining much as at present. Professor Gamow, however, suggests that the sun may become a nova or a supernova in only ten thousand million years, in which case the earth and moon will be melted, and the subsequent history of the solar system will be simply a record of the gradual cooling of lifeless sun and planets. J. H. A.

The Human Meaning of Science, by ARTHUR H. COMPTON. Pp. xii + 88. (Chapel Hill: University of North Carolina Press, 1942.) 6*s*.

In the second of the three lectures in this volume Prof. Compton discusses the abandonment of determinism in modern physics and its bearing on human free will. " As one whose experiments have been partly responsible for this dramatic reversal of the physicist's point of view, I have been specially interested in tracing what the significance of this change may be to human life and thought." Those who have not yet realized the magnitude of the change of outlook will do well to ponder over this simple and outspoken lecture. Human freedom was in direct contradiction to the older system of physical laws ; either it must be an illusion or we must admit that somewhere in the organism a violation of the laws occurred. It is not in direct contradiction to the new quantum mechanics, though, as Compton points out, " this form of mechanics does not necessarily imply freedom ; by itself it implies chance ". Free will is no longer the antithesis of determinism ; it is the antithesis of chance. In its new form the problem of free will presents questions which require careful investigation. The lecture is mainly occupied with the justification of the new outlook, and little space is left to deal with the question, How can volition overcome chance ? I think the difficulty appears more formidable than it really is, because probability in statistical physics is commonly used in a loose way which would not be tolerated by statisticians. More particularly, in the general framework of equations of quantum theory (where there is no reference to special information), the probability is necessarily relative to assumed information which is not explicitly stated. From this point of view the setting aside of chance by volition is no more than the setting aside of *assumed*, but in this case erroneous, information that no volition is involved. Compton has some illuminating remarks on the relation of habit and free will, which may perhaps be summarized by describing habit as statistical volition.

The other lectures are on " Science, Religion and Human Growth " and " The Need for God in an Age of Science ". In the former the main thesis is that, whereas in historic times the human individual has changed little, and it is in the social characteristics of humanity that the chief growth has occurred, science is foremost in accelerating the social change. The latter is an appropriate reflection on the significance of the facts brought out in the two previous lectures. A. S. EDDINGTON.